BONUS BIO-C3

Biodiversity changes: causes, consequences and management implications

Deliverable No: 5.1		Workpackage number and leader: 5, Piotr Margonski	
Date:	31.10.2017	Delivery due date: 31.10.2017	Month 46
Title:	Response of biodiversity indicators to management measures (test of indicators)		
Lead partner for deliverable:		Anastasija Zaiko, Klaipeda University (P8, KU)	
Other contributing partners		Helmholtz Centre for Ocean Research, Kiel (GEOMAR) (P1), Technical University of Denmark (P2), Stockholm University (P4), National Marine Fisheries Research Institute (P5), University of Tartu (P6), Finish Environmental Institute (P7), Thünen Institute of Baltic Sea Fisheries (P11)	
Authors		Anastasija Zaiko (lead) Joanna Calkiewicz, Margit Eero, Christian von Dorrien, Harri Kuosa, Riina Klais, Maiju Lehtiniemi, Piotr Margonski, Daniel Oesterwind, Henn Ojaveer, Andrea Rau, Thorsten Reusch, Anna Törnroos, Jan Warzocha, Monika Winder	
Zaiko A., Calkiewicz J., Eero M., von Dorrien C., Kuosa H., Klais R., Lehtiniemi M., Margonski P., Oesterwind D., Ojaveer H., Rau A., Reusch T., Törnroos A., Warzocha J., Winder M. (2017) *Response of biodiversity indicators to management measures (test of indicators)* BIO-C3 Deliverable, D5.1. EU BONUS BIO-C3, 50 pp., with Appendices. DOI: 10.3289/BIO-C3_D5.1.			
Dissemination level (PU=public, PP=Restricted, CO=confidential)			PU
Nature of the Deliverable (RE=Report, OT=Other)			RE

Acknowledgements

The research leading to these results is part of the BIO-C3 project and has received funding from BONUS, the joint Baltic Sea research and development programme (Art 185), funded jointly from the European Union's Seventh Programme for research, technological development and demonstration and from national funding institutions.

BIO-C3 overview

The importance of biodiversity for ecosystems on land has long been acknowledged. In contrast, its role for marine ecosystems has gained less research attention. The overarching aim of BIO-C3 is to address biodiversity changes, their causes, consequences and possible management implications for the Baltic Sea. Scientists from 7 European countries and 13 partner institutes are involved. Project coordinator is the GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany, assisted by DTU Aqua, National Institute of Aquatic Resources, Technical University of Denmark.

Why is Biodiversity important?

An estimated 130 animal and plant species go extinct every day. In 1992 the United Nations tried countering this process with the "Biodiversity Convention". It labeled biodiversity as worthy of preservation – at land as well as at sea. Biological variety should not only be preserved for ethical reasons: It also fulfils key ecosystem functions and provides ecosystem services. In the sea this includes healthy fish stocks, clear water without algal blooms but also the absorption of nutrients from agriculture.

Biodiversity and BIO-C3

To assess the role of biodiversity in marine ecosystems, BIO-C3 uses a natural laboratory: the Baltic Sea. The Baltic is perfectly suited since its species composition is very young, with current salt level persisting for only a few thousand years. It is also relatively species poor, and extinctions of residents or invasions of new species is therefore expected to have a more dramatic effect compared to species rich and presumably more stable ecosystems.

Moreover, human impacts on the Baltic ecosystem are larger than in most other sea regions, as this marginal sea is surrounded by densely populated areas. A further BIO-C3 focus is to predict and assess future anthropogenic impacts such as fishing and eutrophication, as well as changes related to global (climate) change using a suite of models.

If talking about biological variety, it is important to consider genetic diversity as well, a largely neglected issue. A central question is whether important organisms such as zooplankton and fish can cope or even adapt on contemporary time scales to changed environmental conditions anticipated under different global change scenarios.

BIO-C3 aims to increase understanding of both temporal changes in biodiversity - on all levels from genetic diversity to ecosystem composition - and of the environmental and anthropogenic pressures driving this change. For this purpose, we are able to exploit numerous long term data sets available from the project partners, including on fish stocks, plankton and benthos organisms as well as abiotic environmental conditions. Data series are extended and expanded through a network of Baltic cruises with the research vessels linked to the consortium, and complemented by extensive experimental, laboratory, and modeling work.

From science to management

The ultimate BIO-C3 goal is to use understanding of what happened in the past to predict what will happen in the future, under different climate projections and management scenarios: essential information for resource managers and politicians to decide on the course of actions to maintain and improve the biodiversity status of the Baltic Sea for future generations.

Table of Contents

I. Executive Summary

Identification and quantification of pressure-state links of biodiversity indicators is one of the major scientific challenges ahead. This task aimed at identifying and quantifying the pressure-state links of selected biodiversity indicators, their relevance and response to management measures. Relevant indicators have been selected based on their relevance to: (i) the important anthropogenic pressures in the Baltic Sea and the climate change; (ii) major biodiversity components (i.e. phyto- and zooplankton, benthic invertebrates, fish and NIS); (iii) HELCOM policies, current trends in marine research and methodological developments. The selected indicators/groups of indicators included:

- *Predator fish indicators*
- *NIS indicators*
- *Benthic inverstebrate indicators*
- *Zooplankton indicators*
- *Food web and phytoplankton indicators*
- *Trait-based and functional diversity indicators*
- *Metabarcoding-based indicators*
- *Genetic diversity indicators*

The selection was based on the analysis of the comprehensive catalogue of biodiversity indicators, developed in collaboration with the recent DEVOTES project. In this catalogue, over 600 indicators were compiled, which were developed and used in the framework of different initiatives (e.g., EU policies, research projects) and in national and international contexts (e.g., Regional Seas Conventions, and assessments in non-European seas). Out of these, nearly 200 indicators were reported for the Baltic Sea, representing different pressures, DPSIR stages and target biodiversity groups.

The selected indicators were tested for their performance (incl. stability and sensitivity), based on the outcomes of the Bio-C3 case studies, modelling results and literature reviews, as well as high-level expert knowledge represented by the project consortium. Along with the generalized summary, Deliverable 5.1 provides practical observations, recommendations and precautions for using the existing and newly developed biodiversity indicators.

The summarized management advice resulting from this task includes (but not limited to):

- *Wider employment of ecosystem-based approaches in monitoring and environmental status assessment*
- *Development of adaptive and flexible management frameworks*
- *Continued and consistent monitoring of all biodiversity components across Baltic sub-regions*
- *Uptake and further development of emerging molecular methods for routine monitoring*
- *Switching from simplistic biodiversity metrics towards complex, function-focused multi-trophic indicators, involving tiered assessment approaches*
- *Work toward filling the existing gaps in knowledge on the synergetic effects of multiple pressures and relevant biodiversity response*

More detailed information on the considered indicators/groups of indicators is presented in the form of individual factsheet within the Core Activity section, as well as numerous publication resulting from the Bio-C3 activities and referred to in the text of the Deliverable 5.1.

II. Introduction

A number of recent policy initiatives aimed at biodiversity conservation and overall improvement of marine environmental status, require robust and reliable biodiversity indicators for assessing the current state, clearly linked to pressures and able of informing managers on necessity of response measures.

Thus, a primary goal of the Baltic Sea Action Plan (HELCOM 2007), the EU Marine Strategy Framework Directive and the Convention of Biological Diversity is to reach a favourable conservation status of biodiversity. Biodiversity is influenced by a variety of human activities on land and at sea, as well as by climate change, ecological interactions and conservation measures, which calls for systematic analyses and robust indicators to assess the status of biodiversity. Much of the indicator development so far has focused on the ecosystem state, while establishing links between state and pressure largely remains a future challenge (Bergström et al. 2010, Piet et al. 2010). Moreover, the vast number of existing and continuously emerging biodiversity indicators makes the selection of a fit-to-purpose suite of indicators a non-trivial task.

The properties of the indicators relevant for environmental status assessment and implementing adaptive management strategies have been widely discussed in recent years. Thus, according to different authors, to be considered appropriate for ecological status assessment an indicator should meet the following criteria: be scientifically based; ecosystem relevant and biologically important; responsive, sensitive, specific and predictable; accurate and practical in terms of measurability and cost effectiveness (see e.g. Rice and Rochet 2005, Niemeijer and de Groot 2008, Kershner et al. 2011, Elliott 2011).

The success of management is partially dependent on the availability of scientific tools to managers (Rist et al. 2013; Knights et al. 2014). Robust indicator selection, transparent use of information, and effective communication of results, awareness of potential caveats and emerging improvement opportunities constitute crucial parts of this process. Therefore, the main aim of this task (Deliverable 5.1) was to deliver a tangible advice for stakeholders on relevant data and monitoring needs for robust biodiversity assessment, recommend biodiversity indicators and candidates for targets and threshold values, that will contribute to the development of the evaluation framework for holistic management of the Baltic Sea ecosystem (Bio-C3 Deliverable D5.3).

To achieve the goal, we were following the so-called "funnel approach" for selecting and assessing the most relevant biodiversity indicators for addressing environmental management needs and based on their performance, actual and potential response to management activities (Fig. 1). The core activity of the current task has resulted in a series of comprehensive overviews on prioritised indicators (or group of indicators), scrutinized by the consortium in the course of case studies, modelling exercises and comprehensive literature reviews conducted within WPs 1, 2, 3 and 4 of the Bio-C3 project. Each of these overviews, besides from the general information on the current state of an indicator, its link to a particular pressure (see also Bio-C3 Deliverable D3.1), data needs and potential constrains, contains information on the further improvement opportunities and summarized management advice, that can be easily translated to the stakeholders and decision-makers.

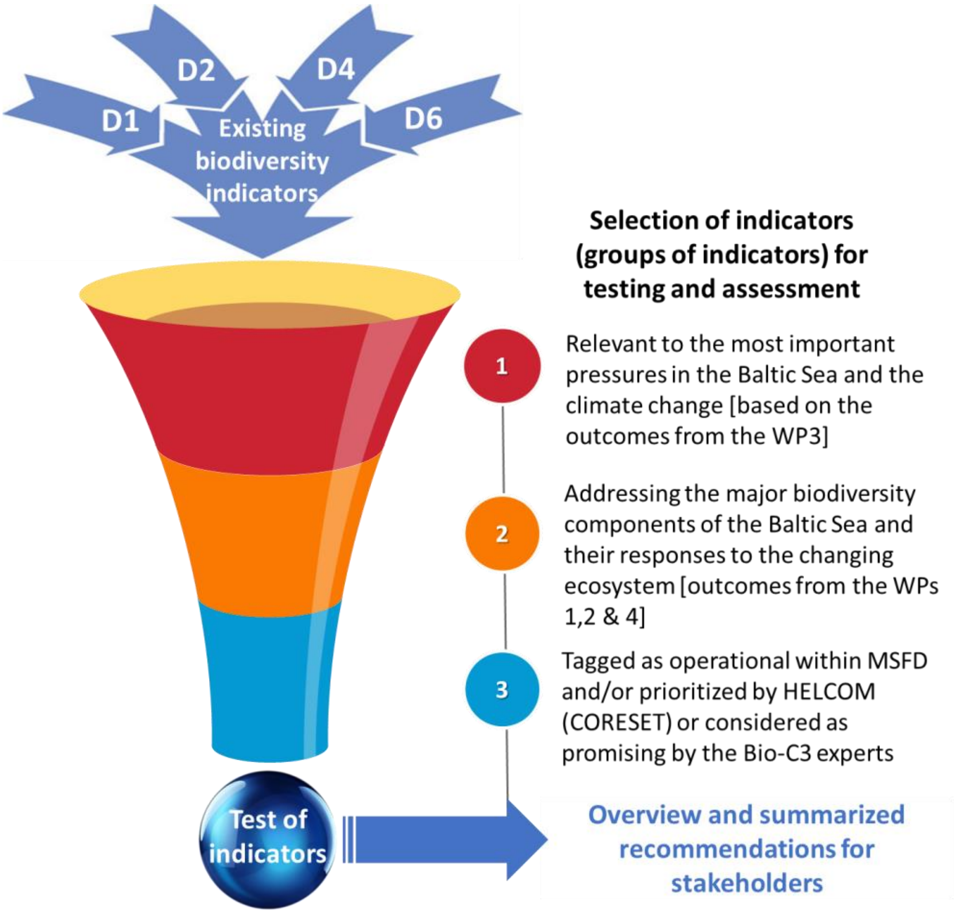


Figure 1. Conceptual scheme of the "funnel approach" for selecting and prioritising indicators for testing within the Task 5.1.

Catalogue of marine biodiversity indicators and prioritized indicators. ***Anastasija Zaiko (P8)***
In collaboration with DEVOTES project (www.devotes-project.eu), a Catalogue of Marine Biodiversity indicators was developed with the aim of providing the basis for assessing the environmental status of the marine ecosystems. This catalogue provides a comprehensive collection of indicators (and relevant metadata) capable of supporting the assessment of four descriptors of MSFD:

- D1: biological diversity
- D2: non-indigenous species
- D4: food webs
- D6: seafloor integrity

Over 600 indicators were compiled, which were developed and used in the framework of different initiatives (e.g., EU policies, research projects) and in national and international contexts (e.g., Regional Seas Conventions, and assessments in non-European seas). The catalogue is freely available through the DEVOTool software application, which provides browsing and query options for the associated metadata Version 8 of the catalogue will be available shortly. The information on this catalogue was summarized in the joint DEVOTES and BIO-C3 publication (Teixeira et al. 2016). The available indicators were reviewed according to their typology, data requirements, development status, geographical coverage, relevance to habitats or biodiversity components, and related human pressures. This initiative has helped establishing the baseline for indicator selection exercise within the Task 5.1 of the Bio-C3 project.

Thus, general analysis of the catalogue allowed identifying the main attributes of the indicators, their geographical coverage and potential for addressing relevant pressures. The catalogue currently contains 611 indicators (catalogue v7), with ~50% of them operational (i.e. tested and validated, with associated target values or classification boundaries, useful for management implications); 36% are under development (indicator is proposed, but for example has not been validated or in the process of calibration). However many of the indicators reported as ”operational” do not have specific set targets, boundaries or reference levels. In regards to the DPSIR stage, only 9% of the recorded indicators are pressure indicators, directly focusing on anthropogenic activities. In terms of addressing pressures, nutrient and organic matter enrichment (eutrophication) was the one addressed by majority of indictors, followed by physical damage and loss of marine habitats.

Interestingly, despite the low biodiversity, a comparatively high number of indicators (196) have been reported from the Baltic Sea (Fig. 2). This reflects overall governmental concern of environmental state in the region, and long-term history of biodiversity research. For the purposed of the Deliverable 5.1, these indicators have been reviewed separately in regards to their typology (DPSIR relevance, see e.g. Oesterwind et al. 2016), data requirements, development status, relevance to biodiversity components, and related human pressures.

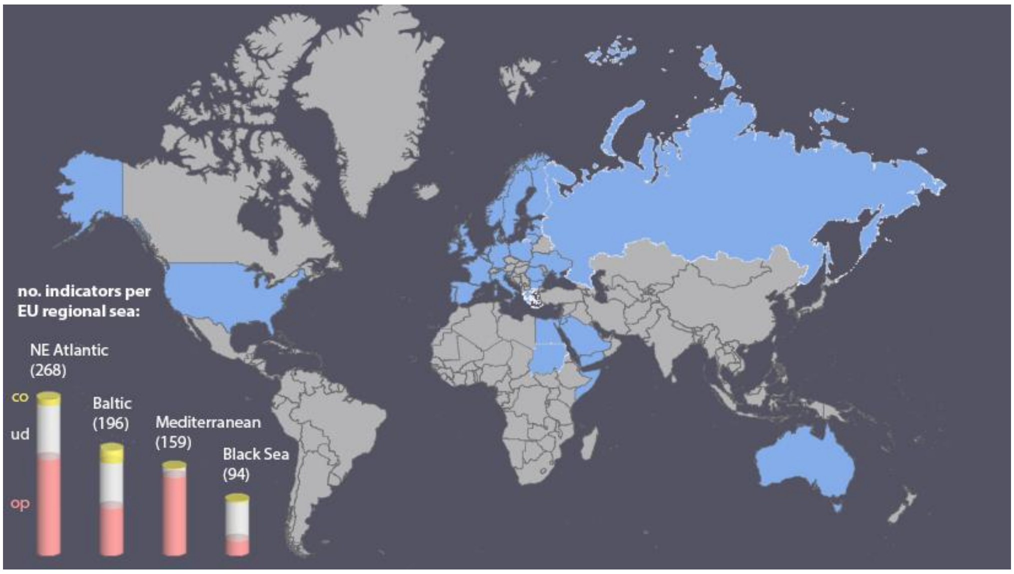


Figure 2. Number of indicators available per European regional sea (either operational–op, under development–ud, or conceptual–co, Teixeira et al. 2016).

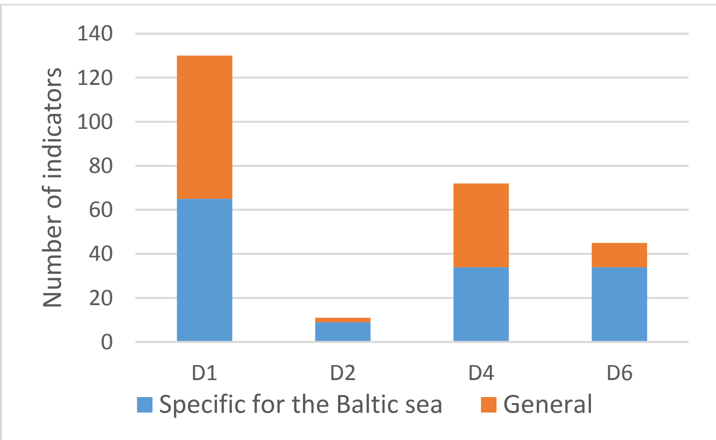


Figure 3. Relevance of the Baltic Sea indicators to MSFD descriptors (note: an indicator can be relevant to one or more MSFD descriptors).

Thus, many of the considered indicators are specific for the Baltic Sea (i.e. specifically developed or modified for application in this ecosystem), with absolute majority addressing biodiversity descriptor

D1 of the MSFD (Fig. 3). In general, the existing indicators (Fig. 4) adequately represent all major biodiversity components of the Baltic Sea.

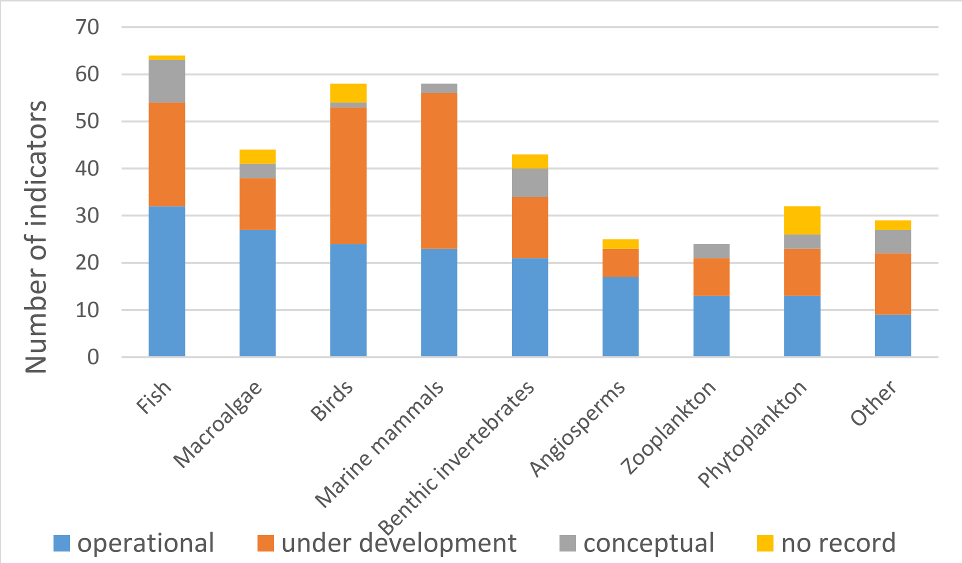


Figure 4. Representation of different biodiversity components by the Baltic Sea indicators (note: an indicator can be relevant to one or more biodiversity components).

Comparing to other regional seas considered in the catalogue, the Baltic Sea has higher number of the pressure indicators (15%, Fig. 5). Most of them are focusing on the specific target groups (e.g. non-indigenous species) and anthropogenic activities (e.g., fish or benthic invertebrates).

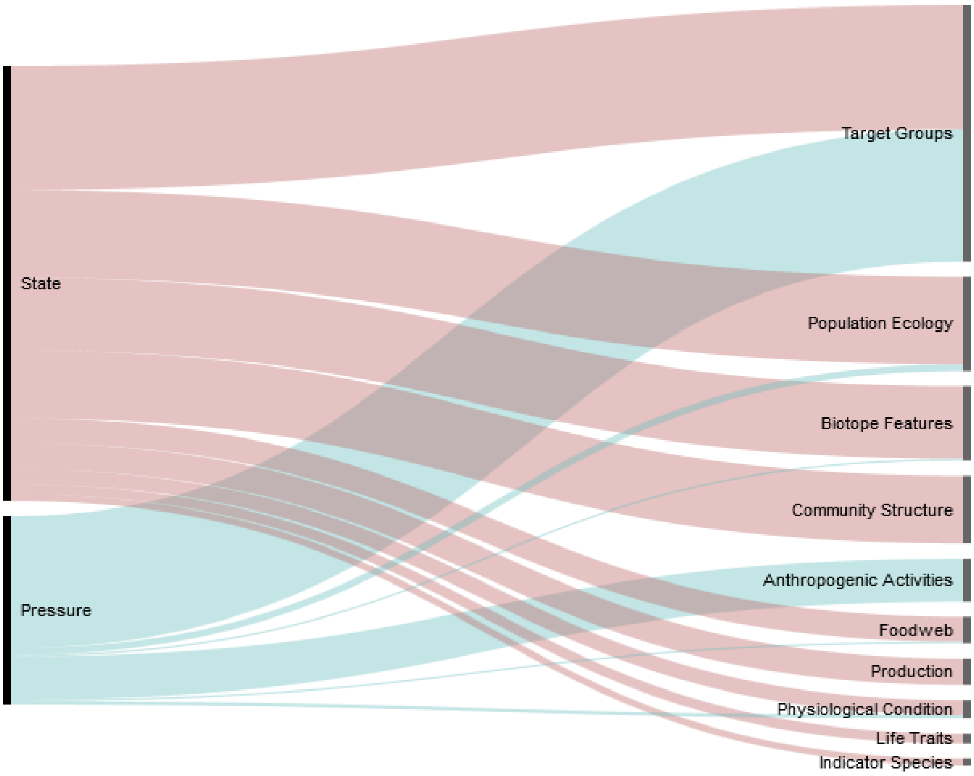


Figure 5. DPSIR stage of the Baltic Sea indicators versus main attributes addressed.

In terms of addressing the most important Baltic Sea pressures (see Bio-C3 Deliverable D3.1), the existing suite of the Baltic Sea indicators compiled in the catalogue gathers a fair diversity of approaches (Fig. 6). Many indicators address nutrient and organic matter enrichment, which reflect eutrophication that is still the most widespread pressure in marine and coastal waters in Europe (EEA, 2013, 2015). Other pressures that were targeted by a few indicators (either directly or indirectly) were related to extraction of species (i.e. fishing), non-indigenous species (NIS), physical loss and physical damage to marine habitats. These reflect the abrasion pressures caused by demersal fishing and aggregate dredging, but also silting, smothering, and increase of turbidity due to coastal and underwater constructions (e.g., Oesterwind et al. 2016). A third group of indicators reflect the effects caused by interference with hydrodynamic processes contamination and extraction (i.e. by-catch and other species removal) pressures. Pressures that have been identified recently such as marine noise, litter or acidification are represented by few indicators.

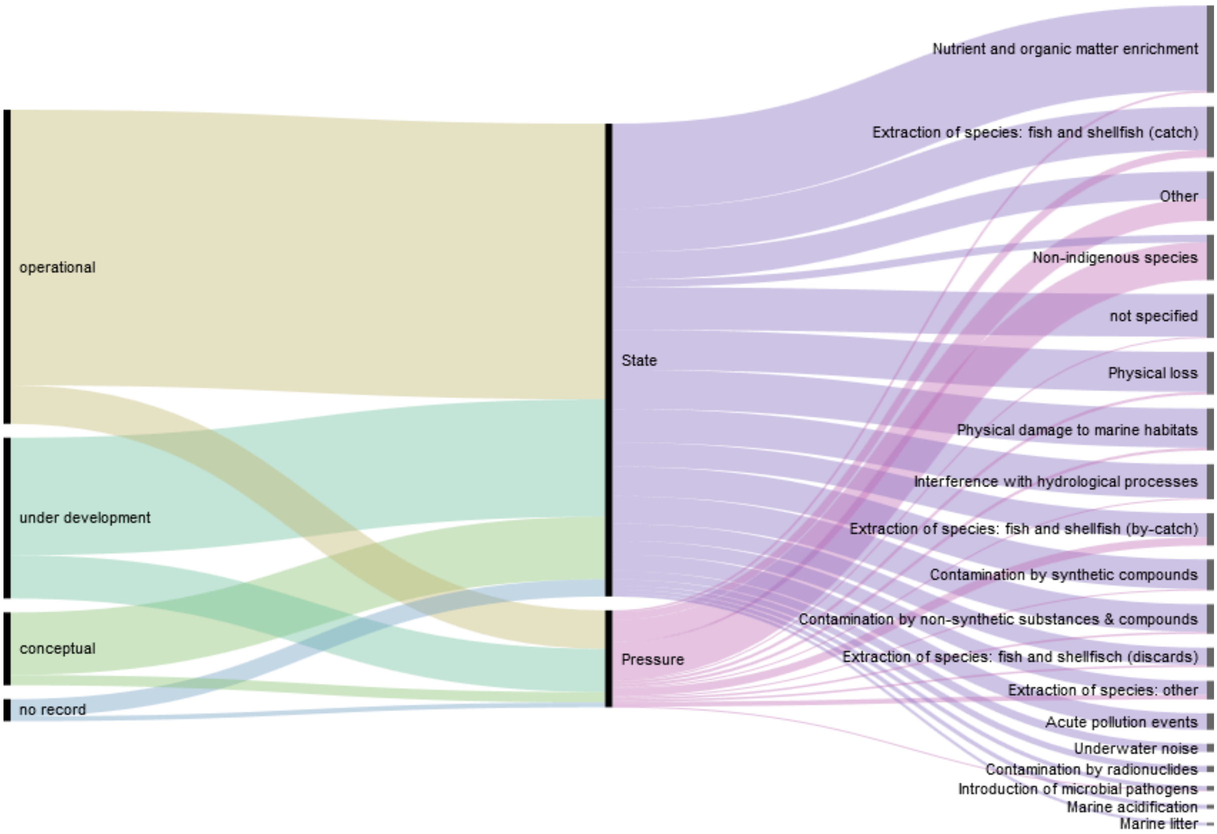


Figure 6. Baltic Sea pressures covered by existing indicators.

Similar patterns of the pressures addressed can be reported for the subset of the Baltic Sea indicators considered by HELCOM (34) as well as those tagged as operational in this subset (22, Fig. 7), rather consistently representing the pressures of the major concern in the Baltic Sea identified in Deliverable 3.1. Thus, based on the results of the catalogue analysis and following the indicator selection approach described above (Fig. 1), the following indicators/groups of indicators were prioritised for consideration in Task 5.1 and testing their performance:

- Predator fish indicators
- NIS indicators
- Benthic invertebrate indicators
- Zooplankton indicators
- Food web and phytoplankton indicators
- Trait-based and functional diversity indicators
- Metabarcoding-based indicators
- Genetic diversity indicators

The results of the Core Activity (test of indicators) is presented in the form of individual factsheets, that summarise the findings derived in the course of the Bio-C3 project, as well as advice for the stakeholders.

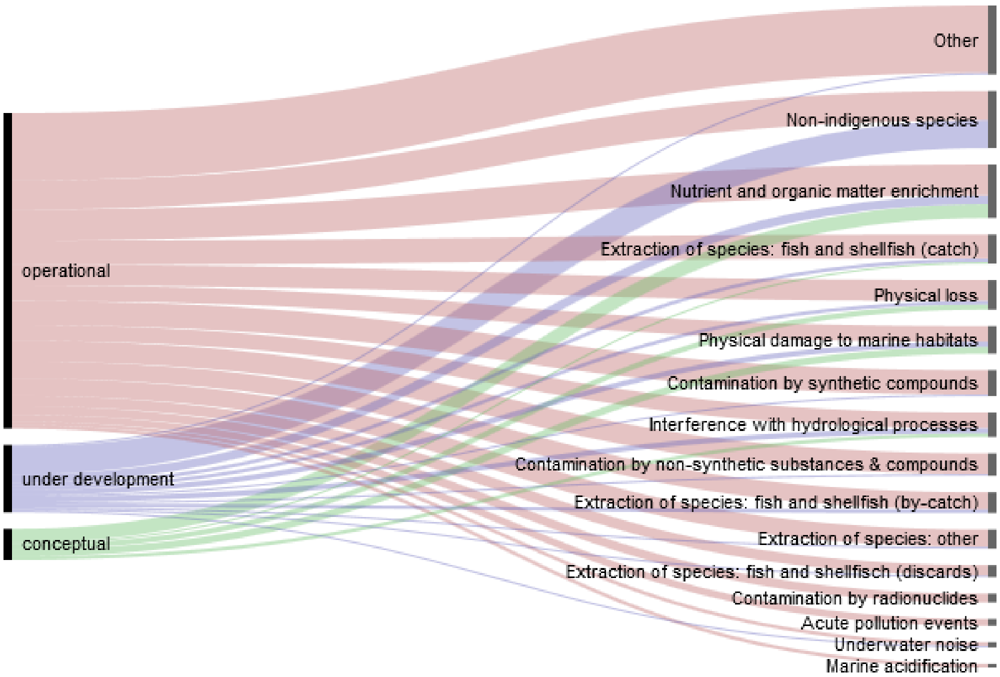


Figure 7. HELCOM indicators *versus* pressures.

Structure of the Core Activity outputs. *Anastasija Zaiko (P8)*

Here we provide brief overview of the factsheets provided for selected indicators/group of indicators in the Core Activity section.

Description
In this section, we describe the selected indicator(s) in terms of:

- Biodiversity component(s) and ecosystem attribute(s) it represents – e.g. fish, benthic inverstebrates, zoo- & phytoplankton, NIS

- Relevance to HELCOM (or other policies)

- Brief scientific background (why it is suggested as an indicator), including reference to existing data and publications, information on the current development status (operational, in development, concept)

- Geographical applicability and coverage in the Baltic Sea (e.g. which sub-region it can be applied for, if there are any spatial restrictions for application, etc.)

- DPSIR stage – pressure or state indicator

- Which pressures, if any, is the indicator related to, either conceptually or practically; if there empirical evidence of the links to particular pressures (strength, direction of change, etc.)

- If a GEnS boundary (for a state indicator) or an environmental target value (for a pressure indicator) has been defined, it is documented here, including the spatio-temporal context of the established boundaries – are they applicable for the entire Baltic, specific sub-regions, habitats, etc.; whether baselines (reference values) are known.

Requirements for deriving the indicator
This is the technical description of how to calculate the indicator, if there are several ways – what is the most practical way, suitable for the Baltic Sea context. Also, monitoring requirements (e.g., where

and when to sample, how much to sample, sample frequency, seasonality, spatial coverage etc, technical requirements & gear, taxonomical expertise required); from minimum requirements to ensure robustness of results to the "ideal world" scenario.

Responsiveness to management measures
Here, we describe (based on the Bio-C3 results, other projects outputs and literature information) whether the indicator responds to management measures (there is empirical evidence of this), has a potential to respond (what are prerequisites for that) or unlikely to respond/response is indirect or not clearly understood.

Note: A good indicator should somehow reflect changes in ecosystem components that are caused by a variation in any specified manageable pressure. It must have a high signal to noise ratio. Ideally, relationship between human activity, resulting pressure on the ecological component and indicator behavior should be clearly understood.

Stability of the indicator
Here we describe whether temporal and/or spatial stability of an indicator performance has been tested and summary of the results.

Constrains and concerns
Here we describe the known constrains and limitations of the indicators.

Advantages and Outlook
Here we justify, what is the general added value of implementing this indicator comparing to other indicators currently used, how it can help in filling gaps understanding the environmental status and pressure dynamics. We also address the methodological outlooks and perspective for improving application of this indicator and overcoming the existing constrains.

Summarized management advice
In a few bullet points, we summarise the most important messages for managers on application of the indicator, including any warning information and general advice, derived while implementing the tasks of the Bio-C3 project.

References
This section holds the references for the citations used in the factsheet for the individual indicator/group of indicators.

References

Bergström L., Diekmann R., Flinkman J., Gardmark A., Kornilovs G. 2010. Integrated ecosystem assessments of seven Baltic Sea areas covering the last three decades 302: 1-90.

EEA, 2013. Balancing the future of Europe's coasts: knowledge base for integrated management. EEA Report No 12/2013. European Environment Agency, Copenhagen, Denmark.

EEA, 2015. State of Europe's seas. EEA Report No 2/2015. European Environment Agency, Copenhagen, Denmark.

Elliott M. 2011. Marine science and management means tackling exogenic unmanaged pressures and endogenic managed pressures – a numbered guide. Marine Pollution Bulletin 62:651-655.

Helcom 2007. HELCOM Baltic Sea Action Plan 1-101.

Kershner J., J. F. Samhouri C. A. James, and P. S. Levin. 2011. Selecting indicator portfolios for marine species and food webs: a puget sound case study. PLoS One 6: e25248.

Knights A.M., F. Culhane, S.S. Hussain, K.N. Papadopoulou, G.J. Piet, J. Raakær, S.I. Rogers, L.A. Robinson. 2014. A step-wise process of decision-making under uncertainty when implementing environmental policy. Environmental Science and Policy, 39: 56-64.

Niemeijer D., and R. S. de Groot. 2008. A conceptual framework for selecting environmental indicator sets. Ecological Indicators 8:14-25.

Oesterwind D., Rau A., Zaiko A. 2016. Drivers and pressures – untangling the terms commonly used in marine science and policy. Journal of Environmental Management, 181: 8-15

Piet G., van Overzee H., Pastoors M. 2010. The necessity for response indicators in fisheries management. ICES Journal of Marine Science 67: 559-566.

Rice J. C., and M.-J. Rochet. 2005. A framework for selecting a suite of indicators for fisheries management. ICES Journal of Marine Science 62: 516-527.

Rist L., A. Felton, L. Samuelsson, C. Sandström, O. Rosvall. 2013. A new paradigm for adaptive management. Ecology & society, 18: 63.

Teixeira H., Berg T., Uusitalo L., Fürhaupter K., Heiskanen A.-S., Mazik K., Lynam C.P., Neville S., Rodriguez J.G., Papadopoulou N., Moncheva S., Churilova T., Kryvenko O., Krause-Jensen D., Zaiko A., Verissimo H., Pantazi M., Carvalho S., Patricio J., Uyarra M.C., Borja A. 2016. A catalogue of marine biodiversity indicators. Frontiers in Marine Science 3: Article207.

III. Core Activity

1. Predator fish indicators. *Margit Eero (P2), Daniel Oesterwind,* Andrea Rau, Wolfgang Nikolaus Probst and Christian von Dorrien *(P11)*

Description

The central Baltic Sea food web is characterized by few species (HELCOM 2009, Ojaveer et al. 2010), while the demersal fish community is dominated by one single species, the cod (Oesterwind et al. 2013). Cod is the key predator fish in the ecosystem preying on pelagic clupeids sprat and herring, and on benthic invertebrates, which availability in turn affects the status of cod as a predator. The structuring impact of biodiversity on ecosystems and food-webs is to a large extent driven by trophic interactions, as changes in the food web propagate to higher trophic levels and affect the entire ecosystem functioning. Because of their relatively low complexity, Baltic sub-regional ecosystems may be less resilient to perturbations impacting individual species exhibiting important ecosystem functions, such as cod (see BIO-C3 deliverables 2.1, 2.2 and 4.2 for further background). Consequently, cod can be considered as an important indicator species for food web and ecosystem functioning in the central Baltic Sea, by picking up changes in lower trophic levels and in the environment, which are reflected in the status of cod, and which in turn feed back to the lower trophic levels.

The biomass of cod is traditionally used as an indicator of stock status (D3 in MSFD) in the context of fisheries management. In BIO-C3, we are focusing on other indicators of cod status, where pronounced adverse developments have been observed in later decades, potentially related to wider ecosystem changes. These include i) nutritional condition, ii) size at first maturation (Lmat) and iii) size structure of the stock (L95) (Fig. 1.1). There is an increasing focus and interest in these variables as potentially useful indicators: first of all, for the status of cod stock itself (e.g. addressed in BONUS INSPIRE). Moreover, these cod indicators are also considered in the new comprehensive HELCOM overview of the state of the Baltic Sea, as part of evaluation of biodiversity status (http://stateofthebalticsea.helcom.fi/biodiversity-and-its-status/fish/#size-structure-of-fish).

Fish condition has been put forward as a potential food-web indicator in ICES (2014a), and suggested as a useful indicator for marine ecosystem status in the context of MSFD (ICES 2016a), being mainly relevant for D1 and D4. The indicators representing demographic characteristics of a fish stock, such as size structure and maturation, are formally included in D3.3 in MSFD. While there are no indicators with operational thresholds internationally agreed upon yet further research and development in this field are necessary: The size distribution within a stock can provide important additional information to the traditional fish stock assessment (Probst et al. 2016). Given the prominent role of cod in the central Baltic ecosystem, where the presence or absence of larger individuals can have measurable effects on food web and ecosystem functioning, the predator stock structure is relevant in relation to D1 and D4 as well. We also consider Large Fish Indicator (LFI), which is originally developed as a fish community indicator (D4); however in the Baltic Sea it is largely driven by cod (Oesterwind et al. 2013).

In BIO-C3, and in other parallel projects (e.g. BONUS INSPIRE), a number of research activities have addressed aspects of cod stock status, following the unexpected adverse developments in cod biology in later years (Eero et al. 2015). For overview of the drivers and consequences of these changes for the Baltic ecosystem, see BIO-C3 deliverables 2.1, 2.2, 4.2. These findings provide a basis for evaluating whether the cod indicators addressed here could be useful in the management context, to guide management decisions considering wider ecosystem interactions and pressures. Cod nutritional condition in the Baltic Sea can be considered as an indicator for feeding opportunities, both regarding pelagic fish and benthic invertebrates, which availability is influenced by fisheries but also by eutrophication and climate affecting environmental conditions, such as oxygen (Eero et al. 2012; Casini et al. 2016a). Nutritional condition can also indicate natural mortality of this predator species due to potential starvation in the periods of low condition (Casini et al. 2016b). In addition to food availability, several other factors have been suggested to have contributed to the observed decline in cod condition since the 1990s. These include direct effects of hypoxia (Casini et al. 2016a), which is related to eutrophication and climate drivers, as well as parasite infestation due to high abundance of grey seals (Horbowy et al. 2016). Further, selective fishing, which truncates size structure in the stock could have caused density dependence and thereby reduced condition (Svedäng and Hornborg 2014). The relative abundance of larger individuals (Orio et al. 2017) and size at maturation (Köster et al. 2017) have reduced nearly simultaneously with the declining condition since the 1990s (Fig. 1.1). Truncated size structure of the cod stock can be related to growth and/or mortality, both of which can potentially be affected by combinations of the same pressures as condition, which could as well be involved in reduced size at maturation, though the mechanisms are not well understood. Further, the cod indicators addressed here can possibly impact one another, e.g. lack of large individuals in the stock can affect condition and vice versa, and both of these indicators can as well be connected to size at maturation. Therefore, these variables are considered here together as indicators for the state of the key predator species.

GES boundaries for these indicators are not formally set. Among the indicators considered here (i.e. condition, Lmat, L95, LFI), the thresholds may be relatively more straightforward regarding condition, where below certain thresholds increased mortality or severe consequences for reproduction occur (Marteinsdottir and Begg, 2000; Dutil and Lambert 2000). Although such experiments have not been conducted for the Baltic cod, and the thresholds in the Baltic may not be the same as found for other stocks in other areas. One of the possible approaches for evaluating the current status in relation to some reference conditions is investigating long-term changes in an indicator, which is facilitated through long research history and data collection for the Baltic cod. In BIO-C3, we have compiled data on selected cod indicators as far back in time as possible, to investigate their dynamics under various combinations of potential pressures to evaluate the use on these indicators in the management context.

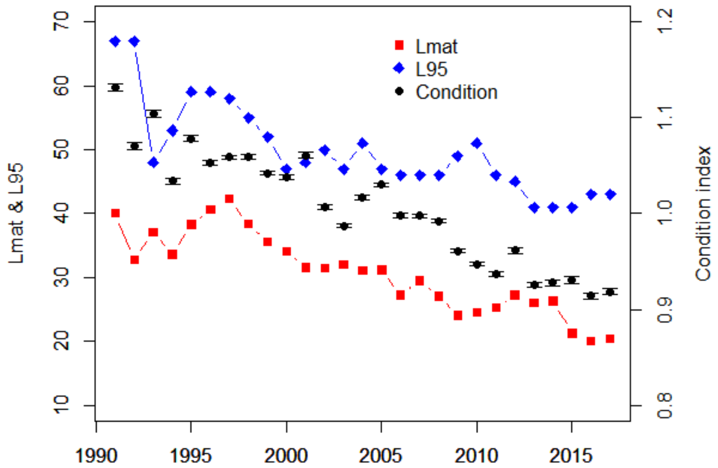


Figure 1.1. Time series of condition (Fulton's K) of 40-60 cm cod, size at first maturation (Lmat) and length corresponding to 95-percentile in the length distribution (L95). Based on data from International Bottom Trawl Survey in Q1 (ICES DATRAS database).

Requirements for deriving the indicator

The data for the indicators considered here (condition, Lmat, L95, LFI) can be obtained from BITS (International Bottom Trawl Surveys) conducted in quarters 1 and 4, and the data are available in ICES DATRAS database from 1991 onwards. For earlier years, used in the analyses in BIO-C3, data have been compiled from literature reports and archival materials of the research institutes.

Condition

Fish condition indices can be calculated in different ways, using whole weight, gutted weight, liver weight, gonad weight etc, in different formulas. The data readily available from DATRAS database for all countries participating in BITS surveys include total length and whole weight measurements. Some countries additionally record information on gutted weight, liver weight and gonad weight, however these data are not internationally available. Thus, the condition index possible to easily calculate based on the information in DATRAS is Fulton's K, where L is total length and W is whole weight of the fish.

$$K = \frac{W}{L^3} * 100$$

The condition index calculated using whole weight can be influenced by gonad development stage and stomach fullness at the time of capture. The influence of gonad development stage on the time series of condition can especially be an issue if changes in spawning time of the fish occur, as has been observed for the Baltic cod (Köster et al. 2017). The measure of gutted weighed is likely less affected by this variability and condition estimates therefore more robust. However, data on gutted weight are not available internationally, while basing the indicator time series on national data, which may be available in some countries, has the disadvantage of lower sample size and coverage than would be the case for international data. Time series of Fulton's K, using gutted weight from Danish national data has been calculated in BIO-C3 (deliverable 2.1), which shows similar trends as the time series with whole weights from DATRAS, confirming that the major decline in condition indices observed since the 1990s is robust to the type of data used. However, international database combining available information on other variables on cod (such as gutted weight) that could be used for calculating different condition indices, to supplement the information derived from whole weight, could be useful.

Lmat

Lmat is defined as length at which 50% of the cod are mature. The indicator can be calculated separately for females and males, for which the international data from BITS surveys are available in DATRAS database. The declining trend in size at first maturation in the 2000s is similar for both females and males (Köster et al. 2017). The indicator of size at first maturation is dependent on maturity determination of the fish at the time when the research cruises are conducted. For the Baltic cod, BITS survey takes place usually in March, while the cod spawning extends from around May until November (Köster et al. 2017). Thus, maturity determination by visual inspection during BITS survey is March, possibly much ahead of the spawning time of the individual cod, can be rather difficult and associated with uncertainties. Thus, maturity determined at the spawning season could be advantageous for calculating this indicator. However, such data are not internationally available presently. In BIO-C3, the maturation of cod at a small size (around 20 cm) has been confirmed by histological methods. Thus, the major reduction in size at maturation in the 2000s, estimated from BITS Q1 data in DATRAS, is probably robust to the uncertainties associated with visual maturity determination. However, regular validation of the maturity determination with histological methods is recommendable. Also, a database gathering information from existing national surveys at the cod spawning time would be useful for this indicator.

L95 & LFI

Several indicators measuring the size structure of a fish population have been proposed (e.g. ICES 2014b, 2016b). In this report, we are using the indicator L95, which has also been considered in ICES as a potentially useful size-based indicator (ICES 2016b). L95 is the length corresponding to 95th percentile in length distribution, and the calculation of this indicator is straightforward.

The L95 indicator in this report is based on catch per unit of effort at length data, which are readily available from DATRAS database from 1991 onwards. For earlier years, length distributions of commercial catches from literature reports and from ICES Working Group materials are used in the analyses conducted in BIO-C3. For recent period, data on length distribution of commercial catches have been internationally compiled from 2000s onwards. As larger cod can possibly be poorly caught in a small-meshed survey gear, supplementing the survey information with commercial data may be beneficial for a representative picture of the status of larger cod in the stock. Thus, compilation of international length distribution data from commercial catches is recommended to be continued in future. The L95 indicator can be sensitive to recruitment (e.g. Probst et al. 2012), which needs to be considered. In BIO-C3, 4 different approaches were tested, including i) full size range ii) only data for >40cm cod, iii) only data for lengths above the peak in length frequency distribution, iv) only data above the length at first catch (50% of mode). The main trends in L95 resulting from the 4 approaches were similar, suggesting that the long-term changes in size structure of cod stock, represented by L95 indicator are robust to the way of dealing with the fraction of smallest cod in the length distribution data. The option ii) was applied in calculation of the L95 indicator, shown in this report.

LFI indicator requires data on all relevant species in the fish community. However, in the Baltic Sea the demersal fish community is largely dominated by cod (Oesterwind et al. 2013). For this reason, the indicator faces major problems in the Baltic, while it is fully operational in the North Sea (Oesterwind et al. 2013). The dominance of cod is especially the case in the eastern Baltic Sea, but also in the western Baltic Sea where flatfishes are relatively more abundant, cod is taking up a large fraction of the demersal fish biomass (Fig. 1.2). This means that only a high proportion of large cod will lead to an LFI being in GES, implying that the aim of the indicator is to achieve a high proportion of large cod. In this case LFI is redundant to MSFD descriptor 3, criterion 3 "Population age and size distribution".

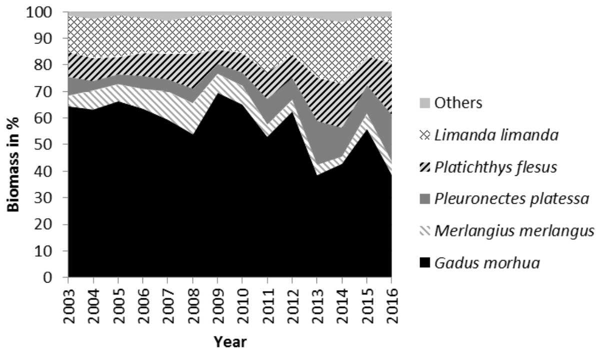


Figure 1.2. Biomass per demersal species during BITS Q1 Survey in the western Baltic (SD-22-24).

Responsiveness to management measures

Concerning indicators of fish stock status, an obvious pressure that can be regulated by management measures is fisheries, and fish biomasses are expected to increase when fishing pressure is reduced. However, a low fishing pressure may not correspond to a good status of those indicators of a fish stock that reflect wider ecosystem characteristics. This is illustrated by Fulton's K condition index for Eastern Baltic cod, where lower values in fact have coincided with relatively low fishing pressure (Fig. 1.3). Similarly, the lowest values in the time series of L95 are observed in recent years, when the overall fishing pressure is estimated to be relatively low (Fig. 1.3). This demonstrates that traditional single stock fisheries management measures affecting overall fishing mortality of a stock are not necessarily improving the status of indicators such as Fulton's K and L95, which confirms that these indicators are reflecting wider ecosystem properties. Fisheries selectivity would likely affect L95, and possibly also condition (Svedäng and Hornborg 2014), however these effects are not quantified due to unresolved relative contribution of individual pressures to the dynamics of these indicators. Dedicated analyses addressing the performance of various fish indicators, including Lmat and L95 concluded that a clear link between these indicators and specific pressures is lacking (Annex 1).

The decline in cod condition in later decades has been suggested to be partly related to availability of sprat and herring as prey (see BIO-C3 deliverables 2.1, 2.2, 4.2). Thus, spatial management measures regulating fisheries for these species could affect cod condition (Eero et al. 2012), taking into account changes in clupeid distribution in later decades, when the majority of clupeids has been out of reach for cod (BIO-C3 deliverable 2.2). However, as there are number of other possible explanations for reduced cod condition (e.g. hypoxia, benthic prey, seals), which relative importance is not fully resolved, the effect of possible fisheries management measures is unclear. The pressures affecting hypoxia and benthic prey (Casini et al. 2016a) include eutrophication, which responds to management with long time lags, as well as climate, which may not be manageable at all, at least in short term. Abundance of grey seals, affecting cod via parasite infestation (Horbowy et al. 2016) could potentially be regulated by hunting or other management measures, though the effect on cod condition is hitherto not quantified.

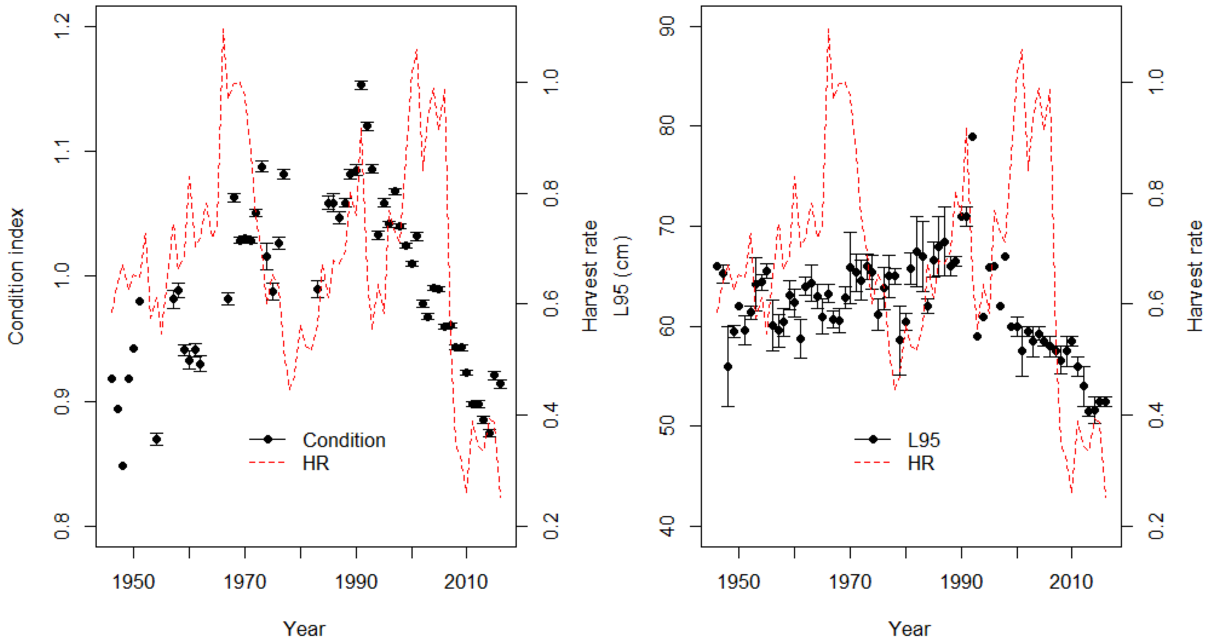


Figure 1.3. Left panel: Fulton's K condition index (mean and standard error of the mean) for 40-60 cm cod in Q1 in SDs 25-26 (for years before 1991, the data includes individual fish measurements from the Baltic Sea Fisheries Institute in Rostock, Polish Marine Fisheries Research Institute, and Technical University of Denmark; after 1991 the data are from international BITS survey. Right panel: L95 as a mean of the available length frequency datasets for a given year, and standard error of the mean. Harvest rate (HR) in both panels is calculated as total catch divided by SSBProxy, which until 2006 is based in ICES stock assessment, and for later years derived from a relationship between survey cpue for >30 cm cod and SSB from stock assessments in years 1991-2006.

Stability of the indicator

Spatially large variations in indicators of cod status have not been detected. For example, the decline in nutritional condition in later decades has been observed in all Subdivisions in the Baltic Sea (Casini et al. 2016a). In a smaller scale, spatial differences may occur, for example in coastal areas versus deep basins, however this has not been sufficiently investigated. Further, there is a large variability between individuals, e.g. in nutritional condition, while the indicators calculated here are measuring developments in population averages.

In a temporal scale, the indicators of predator stock status may at different time periods reflect different aspects of ecosystem conditions and pressures. The 2000s are characterised by negative developments in all cod indicators considered (Fig. 1.1) as well as in most of the ecosystem and environmental parameters suggested as possible explanatory variables (Fig. 1.4). Opposite, the time period with high indicator values in the early 1990s corresponds to good conditions for cod in nearly all aspects (Fig. 1.4). This complicates disentangling the main pressures behind the adverse changes observed in the predator stock in later decades. In BIO-C3, the time series on condition and size structure in the stock have been extended back in time, to elucidate their dynamics under different combinations of pressures. Interestingly, cod condition seems to have been low also in the 1940s-early 1950s, when the combination of ecosystem and environmental conditions, as well as size structure of the stock, has generally been more positive than in the 2000s. This suggests that the mechanisms affecting the condition of the predator stock may be even more complex than hitherto thought, and possibly other factors are involved that have so far not been considered. Further, the entire ecosystem setup and functioning, also from cod perspective, was probably considerably different back in time (Eero et al. 2015), in the period with frequent major Baltic inflows and before the ecosystem regime shift (Möllmann et al. 2009). Thus, the relationships between cod indicators and respective impacting factors may have changed over time.

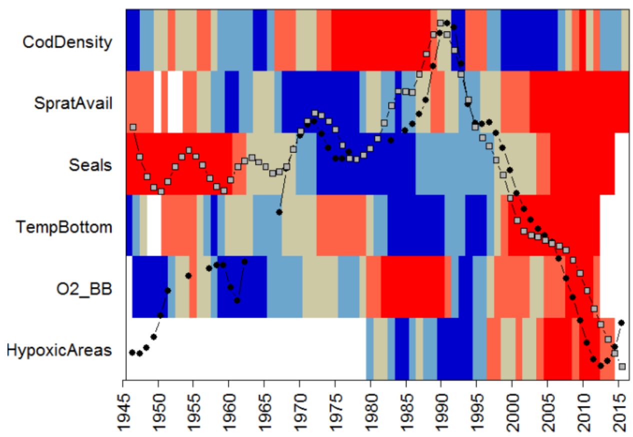


Figure 1.4. Long term relative changes in ecosystem and environmental variables considered to influence cod in the Baltic Sea (the color categories represent 20th percentiles of the range of observed values during the analyzed period, representing beneficial (blue) and negative (red) conditions for cod), overlaid with smoothed trends in Fulton's K (black circles) and L95 (grey rectangles) indicators.

Constrains and concerns

Multiple drivers and pressures affect biodiversity and ecosystem functioning in the Baltic Sea, including its key predator fish, i.e. cod. This limits the use of particular values for cod indicators in direct conjunction with specific management measures. Further, the research conducted in BIO-C3 and elsewhere has demonstrated multi-decadal changes in the Baltic ecosystem driven by large scale processes, such as major Baltic inflows. Thus, new ecosystem setups may emerge, implying that the conditions faced by individual ecosystem components such as cod and the dominant pressures may change over time (e.g. Eero et al. 2015, Köster et al. 2017). This complicates decisions on possible management measures to mitigate the undesired developments in cod and in the ecosystem and requires continuously updated process understanding. In situations where clear link to pressures is not identified or quantified, advice can at best only be given on the required direction of action (Annex 1).

Advantages and Outlook

BIO-C3 and other projects have greatly advanced our understanding of the processes affecting the status of cod as a key predator in the central Baltic ecosystem. Also, new knowledge is gained on the possible consequences of the recent adverse developments in cod indicators for the future of this stock, and thereby for the entire central Baltic ecosystem. Further improvement in understanding of these processes will likely continue in future. The cod indicators addressed here are considered useful as surveillance indicators first of all for the cod stock itself, measuring other aspects of stock condition than only biomass that is regularly monitored, allowing for better predictions of likely developments in future. Moreover, these relatively simple indicators of the predator stock status are useful for detecting major changes in food web and other ecosystem conditions, reflected in changes in the state of cod. This can help to identify new potential management issues that need to be considered, as well as guide forming new relevant research questions under the changing environment of the Baltic Sea.

Summarized management advice

- *Indicators of cod stock status, other than biomass and fishing mortality, are important to consider in fisheries management, e.g. in determining reference values for sustainable exploitation.*

- *Fisheries management measures directed to cod alone do not necessarily improve the status of the indicators of stock condition, but ecosystem based approaches taking into account other pressures and components of the ecosystem are needed.*

- *Indicators of predator stock condition respond to multiple pressures, some of which may not be manageable, at least at short time scales. Thus, an adjustment to a management measure may not result in an immediate improvement in cod indicators.*

- *The changing environmental and ecosystem conditions in the Baltic Sea call for adaptive and flexible management frameworks.*

- *Continued monitoring of the cod indicators combined with updated process understanding is needed to support appropriate management decisions corresponding to a given ecological setup.*

References

Casini M., Käll, F., Hansson, M., Plikshs, M., Baranova, T., Karlsson, O., Lundström, K., Neuenfeldt, S., Gårdmark, A., Hjelm, J. 2016a. Hypoxic areas, density-dependence and food limitation drive the body condition of a heavily exploited marine fish predator. Royal Society Open Science. 3: 160416.

Casini, M., Eero, M., Carlshamre, S. and Lövgren, J. 2016b. Using alternative biological information in stock assessment: condition-corrected natural mortality of Eastern Baltic cod. ICES Journal of Marine Science, 73: 2625-2631. doi:10.1093/icesjms/fsw117.

Dutil, J. D., and Lambert, Y. 2000. Natural mortality from poor condition in Atlantic cod (Gadus morhua). Canadian Journal of Fisheries and Aquatic Sciences, 57: 826–836.

Eero, M., Vinther, M., Haslob, H., Huwer, B., Casini, M., Storr-Paulsen, M., Köster, F.W. 2012. Spatial management of marine resources can enhance the recovery of predators and avoid local depletion of forage fish. Conservation Letters 5: 486-492.

Eero, M., Hjelm, J., Behrens, J.,Buchmann, K., Cardinale, M., Casini, M.,Gasyukov, P., Holmgren, N.,Horbowy, J.,Hüssy, K., Kirkegaard, E., Kornilovs, G., Krumme, U., Köster, F. W., Oeberst, R., Plikshs, M., Radtke, K., Raid, T., Schmidt, J., Tomczak, M. T., Vinther, M., Zimmermann, C., and Storr-Paulsen, M. 2015. Eastern Baltic cod in distress: biological changes and challenges for stock assessment. ICES Journal of Marine Science, 72: 2180–2186.

Helcom. 2009. Biodiversity in the Baltic Sea – An integrated thematic assessment on biodiversity and nature conservation in the Baltic. Balt. Sea Environ. Proc. No. 116B

Horbowy, J., Podolska, M., Nadolna-Ałtyn, K. 2016. Increasing occurrence of anisakid nematodes in the liver of cod (Gadus morhua) from the Baltic Sea: Does infection affect the condition and mortality of fish? Fisheries Research. 179: 98–103.

Köster, F. W., Huwer, B., Hinrichsen, Hans-H., Neumann, V., Makarchouk, A., Eero, M., Dewitz, B. V., Hüssy, K., Tomkiewicz, J., Margonski,P., Temming, A., Hermann, Jens-P., Oesterwind, D., Dierking, J., Kotterba, P. and Plikshs, M. 2017. Eastern Baltic cod recruitment revisited—dynamics and impacting factors. ICES Journal of Marine Science, 74 (1): 3-19.

ICES. 2014a. Report of the Workshop to develop recommendations for potentially useful Food Web Indicators (WKFooWI). ICES CM 2014\ACOM:48

ICES 2014b. Report of the Workshop to draft recommendations for the assessment of Descriptor D3 (WKD3R). ICES CM 2014/ACOM:50

ICES 2016a. Report of the Workshop on Fish condition (WKFICON). ICES CM 2016/SSGIEOM:30

ICES 2016b. Report of the Workshop on Guidance on Development of Operational Methods for the Evaluation of the MSFD Criterion D3.3 (WKIND3.3i). ICES CM 2016/ACOM:44

Marteinsdottir, G., Begg, G.A. 2002. Essential relationships incorporating the influence of age, size and condition on variables required for estimation of reproductive potential in Atlantic cod *Gadus morhua. Mar Ecol Prog Ser,* 235: 235–256

Möllmann C., Diekmann R., Müller-Karulis B., Kornilovs G.,Plikshs M. & Axe P. 2009. Reorganization of a large marine ecosystem due to atmospheric and anthropogenic pressure: a discontinuous regime shift in the Central Baltic Sea. *Global Change Biol.* 15, 1377–1393.

Oesterwind, D., Psuty, I., Pachur, M., Dorrien, C.von, Lejk, A., 2013: Proportion of large fish in the community, HELCOM Core Indicator Report. http://www.helcom.fi/Core%20Indicators/HELCOM-CoreIndicator-Proportion_of_large_fish_in_the_community.pdf

Ojaveer H, Jaanus A, MacKenzie BR, Martin G, Olenin S, et al. 2010. Status of Biodiversity in the Baltic Sea. PLoS ONE 5(9): e12467. doi:10.1371/journal.pone.0012467

Orio, A., Florin, A.-B., Bergström, U., Sics, I., Baranova, T., and Casini, M. 2017. Modelling indices of abundance and size-based indicators of cod and flounder stocks in the Baltic Sea using newly standardized trawl survey data. ICES Journal of Marine Science, doi:10.1093/icesjms/fsx005.

Probst, N.W., Rau, A., Oesterwind, D., (2016). A proposal for restructuring Descriptor 3 of the Marine Strategy Framework Directive (MSFD). Journal of Marine Policy, 74:128 – 135.

Probst, W. N., Stelzenmüller, V., and Fock, H. 2012. Using cross-correlations to assess the relationship between time-lagged pressure and state indicators: an exemplary analysis of North Sea fish population indicators. ICES Journal of Marine Science, 69: 670

Svedäng, H., and Hornborg, S. 2014. Fishing induces density-dependent growth. Nature Communications, 5: 4152.

2. NIS arrival indicators. *Henn Ojaveer (P6), Maiju Lehtiniemi (P7), Anastasija Zaiko (P8)*

Description

Non-indigenous species (NIS) are one of the major external stressors for change in marine ecosystems and the impacts they may cause are often unpredictable. Over 130 NIS and cryptogenic species (CS) have been observed in the Baltic Sea (Ojaveer et al. 2016). The primary pathways responsible for the currently established NIS/CS (59% of all introduced species) are shipping and natural spread from neighbouring areas. Substantial uncertainty in the information on introduction pathways (except for deliberate releases) hampers detailed analyses and makes it very difficult to assess new human-mediated introductions both into and inside the Baltic Sea.

The considered NIS arrival indicator is called as '*the number of new NIS introductions via human activities*'. The indicator takes into account all new NIS findings independent whether they have established self-sustaining populations or not. Thus, the number of arrived NIS and CS evaluates both the successfulness of preventive management as well as the status of the ecosystem by indicating the areas where the level of unpredictable risk is high (Olenin et al. 2016).

The ultimate goal of the NIS arrival indicator is to minimize anthropogenic introductions of NIS to zero (HELCOM 2017). The GES-level is achieved if there are no new introductions of NIS per assessment unit through human activities during a six years assessment period. It is suggested that the assessment unit should be regional sea (i.e. whole Baltic Sea; for justification see below). As a mid-term goal, a decrease in the rate of new introductions should be considered. The evaluation against the GES threshold value is carried out by summing all new introductions into the Baltic Sea over the assessment period. To enable an evaluation of status, the indicator requires a baseline in the form of a list of NIS/CS already present by the end of the previous assessment period (i.e. by the beginning of the new assessment period).

The focus of the indicator is on human-mediated introductions and not secondary spread by natural means (migration, water currents etc.). There are large regional inconsistencies in the assignment of introductions to vectors/pathways due to different levels of information certainty/confidence and information availability in different sub-basins. This is the reason why the indicator should consider only new introductions into the whole Baltic Sea where we have in total better cumulative level of confidence by the introduction vector/pathways.

The majority of the relevant data is accessible in AquaNIS (www.corpi.ku.lt/databases/aquanis/), regularly updated by ICES WGITMO members. The processing required for making an evaluation against the threshold value for the Baltic Sea only requires summing the number of new introductions to the Baltic Sea during the assessment period. At present, the indicator is calculated for the whole sea area due to uneven monitoring and research efforts between sub-regions.

The primary relevant policy documents the indicator addresses is HELCOM Baltic Sea Action Plan (No new NIS introductions via shipping; HELCOM 2007) and EU Marine Strategy Framework Directive (D2: Non-indigenous species introduced by human activities are at levels that do not adversely alter the ecosystem; European Commission 2017).

Requirements for deriving the indicator

The parameter used to evaluate whether the GES-threshold is achieved or not in this indicator is the number of new species introductions by human activities to the Baltic Sea during the assessment period. The indicator is based on AquaNIS data that has been verified by national experts. The indicator results could be significantly improved if dedicated monitoring program for NIS would be launched in all HELCOM countries. Currently dedicated monitoring for NIS is ongoing only in a few countries and evaluations are strongly biased towards better investigated groups (molluscs, crustaceans, fish), almost no information on micro/meiotaxa and pathogens is available for consideration.

Responsiveness to management measures
The NIS indicator is directly linked to human pressures as NIS are transported to the Baltic Sea by intentional or unintentional human activity. However it is not directly linked to any specific pressure (e.g. ballast water or hull fouling mediated introductions, aquaculture activities) and thus success or failure in specific management measure is not easy to see from the indicator result.

Stability of the indicator
Not tested.

Constrains and concerns
Dedicated regular NIS monitoring programs are not in place in most countries. Therefore, the data available is not considered to representatively cover all areas of the Baltic Sea to ensure that all new introductions are detected with similar effort. Thus a zero result for an assessment unit may be a false negative.

Another possible sources of false negatives can be attributed to: (*i*) insufficient sampling effort (e.g., if for economy reasons only minor part of the marine area is covered by surveillance grid); (*ii)* inappropriate or ineffective sampling approaches (e.g., if applying core or grab sampling for inferring biodiversity of mobile benthic crustaceans); (*iii*) *insufficient* taxonomic resolution achieved by conventional identification methods (e.g., for micro/meiotaxa); and (*iv*) misidentification due to the lack of taxonomic expertise (especially for small, cryptic taxa and those at larval stages, (Lehtiniemi et al. 2015).

Advantages and Outlook
This indicator gives a good general overview of the human mediated introductions and allows evaluation of the success of preventive management measures. This is the only NIS indicator at present and it is accepted by HELCOM contracting parties, thus making it possible to assess the status of the sea over all Baltic Sea regions.

The precision and predictive capacity for this indicator can be noticeably improved, if traditional surveillance approaches are coupled with the emerging molecular monitoring techniques (e.g., community metabarcoding and targeted species detection using species-specific molecular assays, see for example Ardura et al. 2015, Zaiko et al. 2015 and below section for more details).

Summarized management advice

- *The indicator is relatively robust to evaluate new NIS introductions at the pan-Baltic level (see HELCOM 2017).*

- *Due to limitations in knowledge, in several cases the exact introduction pathway may remain uncertain*

- *Uneven surveillance effort and current taxonomical constraints make sub-regional comparisons relatively uncertain*

- *Uptake of molecular techniques for routine environmental surveillance and development of well-annotated molecular reference database for the Baltic Sea species would improve overall NIS detection capacity and consistency of the GES assessment across sub-regions*

References

Ardura A., Zaiko A., Martinez J.L., Samuiloviene A., Semenova A., Garcia-Vazquez E. (2015). eDNA and specific primers for early detection of invasive species- a case study on the bivalve Rangia cuneata, currently spreading in Europe. Marine Environmental Research, 112(B): 48-55.

HELCOM 2007. HELCOM Baltic Sea Action Plan. Helsinki, 102 pp.

HELCOM 2017. HELCOM (2017): First version of the 'State of the Baltic Sea' report – June 2017 – to be updated in 2018. Available at: http://stateofthebalticsea.helcom.fi

European Commission 2017. COMMISSION DIRECTIVE (EU) 2017/845 of 17 May 2017 amending Directive 2008/56/EC of the European Parliament and of the Council as regards the indicative lists of elements to be taken into account for the preparation of marine strategies. Official Journal of the European Union L125/27-33.

Lehtiniemi M., Ojaveer H., David M., Galil B.S., Gollasch S., McKenzie S., Occhipinti-Ambrogi A., Olenin S., Pederson J. (2015) Dose of truth - Monitoring marine non-indigenous species to serve legislative requirements. Marine Policy 54: 26-35.

Ojaveer, H., S. Olenin, A. Narscius, A.-B. Florin, E. Ezhova, S. Gollasch, K.R. Jensen, M. Lehtiniemi, D. Minchin, M. Normant-Saremba & S. Strake (2016). Dynamics of biological invasions and pathways over time: a case study of a temperate coastal sea. Biological Invasions 19:799-813

Olenin, S., A. Narščius, S. Gollasch, M. Lehtiniemi, A. Marchini, D. Minchin & G. Srėbalienė (2016). New Arrivals: An Indicator for Non-indigenous Species Introductions at Different Geographical Scales. Frontiers in Marine Science 3: 208 DOI:10.3389/fmars.2016.00208

Zaiko A., Samulioviene A., Ardura A., Garcia-Vazquez E. (2015) Metabarcoding approach for non-indigenous species surveillance in marine coastal waters. Marine Pollution Bulletin, 100: 53-59.

3. Metabarcoding-based indicators. *Anastasija Zaiko (P8)*

Description

Baseline biodiversity information in the form of measureable parameters (metrics) is required for deriving indicators on different biodiversity component and ecosystem attributes (e.g. benthos, zoo- and phytoplankton, non-indigenous species (NIS)). The metrics that are generally required can be placed into the following categories: (*i*) general species lists; (*ii*) presence/absence of target species; (*iii*) viability/condition of detected taxa; (*iv*) quantitative data on the abundance of species.

Molecular methods are increasingly being touted as promising tools for addressing existing knowledge gaps and assessing biodiversity for improved environmental management (e.g., Ji et al. 2013; Aylagas et al. 2014; Wood et al. 2013), and potentially can be applied for deriving biodiversity information on different components of marine ecosystems, addressing both state and pressure indicators. But depending on the question addressed each has merits and limitations (see Annex 2 for more details).

Although applicability of metabarcoding-based indicators for environmental monitoring is increasingly addressed in the literature (Aylagas et al. 2014; Pochon et al. 2016; Laroche et al. 2017), they are not yet routinely applied in national monitoring programmes, neither have established GEnS or target values for the Baltic Sea ecosystem.

Requirements for deriving the indicator

Currently there are no well-established or standardized sampling protocols or analytical guidelines for deriving metabarcoding-based indicators. All empirical studies demonstrating their potential use are of rather local scale, not allowing for cross-ecosystem or cross-regional methodological comparisons and standartisations.

Due to high sensitivity of the method, the minimum requirements to ensure robustness of results should include precautionary measures to eliminate contamination risks at each step of sample processing, including incorporation of blank controls.

Using metrics based on multi-marker metabarcoding, reasonable field and lab replication, careful consideration of appropriate sampling methods for the targeted taxa/ecological group (taking into account the biology of their different life stages), and using a reliable reference databases is highly recommended.

In biodiversity assessments, quantitative measurements are generally preferred over qualitative and categorical measurements. In this sense, molecular methods have some limitations as they are challenged when providing estimates of true abundance and/or biomass information (Yu et al. 2012). Semi-quantitative biological information can be derived from molecular studies, e.g. relative quantities of species in the community can be inferred from the percentage of sequence reads obtained through metabarcoding (e.g., Hajibabaei et al. 2011; Evans et al. 2016). However, quantitative biodiversity information derived from environmental DNA samples should be treated differentially for micro- and macro-communities, and for each specific metabarcoding-based metric a well-designed calibration study involving conventional taxonomic assessment is required.

Responsiveness to management measures
Recent elaboration of metabarcoding-based indices have demonstrated the potential for using HTS data to measure diversity change of marine benthic assemblages along pollution gradients (Pochon et al. 2015; Lejzerowicz et al. 2015; Keeley et al. 2017). A generalized conclusion can be made, that metabarcoding-based indices are way more reactive to environmental changes than conventional biodiversity metrics, therefore their response to management measures is expected to be high as well. Being able to detect organisms across all biological domains, metaborcoding can help to derive both signal of change in relation to the all integrated pressures, as well as specific ones (e.g. based on responses of specific taxa or ecological groups, more sensitive to particular stressors).

It should be kept in mind, however, that eDNA metabarcoding allows for detecting free-floating DNA and DNA from non-living cells, and this signal can be detected at a location for quite a while after organism has been present there. This signal can potentially mask the short-term pressure or management-related response. To address this issue, eRNA metabarcoding can be applied for detecting biologically-active biodiversity (as eRNA degrades more rapidly in the marine environment).

Stability of the indicator
Molecular methods offer the advantage of being able to rapidly characterize comprehensive biodiversity from small sample size, and can be effectively applied to a wide range of environments and spatial scales (Lance, Carr 2012; Zaiko et al. 2015; Evans et al. 2016; Shaw et al. 2016). However, their temporal-spatial stability has not been comprehensively assessed in the Baltic Sea. It can be assumed though, that due to high taxonomic resolution afforded by metabarcoding, the variability of detected biodiversity will be higher comparing to conventional methods. Higher stability can be expected for benthic samples, both due to more stable matrix and higher possibility to obtain legacy DNA (extracellular or non-living material, better accumulated and preserved in seabed sediments).

Constrains and concerns
The biggest current limitation of metabarcoding application for environmental monitoring – is incompleteness of the reference sequence databases (Ratnasingham and Hebert, Zaiko et al. 2015), this restricts robust taxonomic assignment of all taxonomic groups from an environmental sample.

Handling of ever-increasing volumes of molecular data requires increased computational resources and analytical efforts. This affects the overall cost structure of a biodiversity research project, requiring a larger budget allocated to the analysis component, comparing to traditional research where most of the cost is spent on experimental work and data generation (Sboner et al. 2011).

Detecting signal from legacy DNA may lead to the detection of false positive signals, e.g. from non-viable organisms, untargeted sources, incidental contamination, or as a result of marker-related biases or imprecise taxonomic assignments (Ficetola et al. 2015).

There are many factors influencing biological biases when estimating organisms' abundance from eDNA samples, including complex intra- and inter-individual biological variation due the presence of different number of nuclei, gene copies, genome sizes and/or biovolume (Zaiko et al. in prep; Annex 2). This impedes deriving robust quantitative biodiversity information from metabarcoding data.

Advantages and Outlook

In general, when compared to the traditional morphology-based biodiversity assessment, the estimated cost per sample applying molecular methods is significantly less due to the constantly increasing high-throughput and multiplexing capacities (Ji et al. 2013).

Sensitivity and resolution of metobarcoding allows early warning of undesirable ecosystem change, well before it can be detected by conventional monitoring techniques (Annex 2).

Metabarcoding can help detecting cryptic marine species, or those sparsely distributed or at early life stage, which identification can be difficult or impossible using microscopy. This is particularly important for NIS managements and timely response to new arrivals (Annex 2).

Ability to characterize comprehensive biodiversity from a small environmental sample using molecular approaches enables better research and consequently management opportunities in vulnerable habitats and ecosystems (Annex 2).

Summarized management advice

- *Metabarcoding techniques have a great potential for deriving biodiversity information for better management of the Baltic Sea ecosystem*

- *Molecular techniques are advancing rapidly and it is likely that extensive scientific effort in this field will overcome many of the current caveats resulting in more robust and cost efficient methods for deriving metabarcoding-based biodiversity indicators.*

- *A dedicated scientific effort is needed to develop fit-for-purpose metabarcoding-based indicators addressing specific biodiversity components, ecosystem attributes and pressures in the Baltic Sea.*

- *To facilitate and encourage effective uptake of metabarcoding-based indicators, there is a need for an international collaborative framework aimed at unifying molecular sample processing and analysis methods for monitoring applications.*

- *Ideally, a comprehensive, region-specific, well-annotated database covering all levels of the Baltic Sea biodiversity and a range of marker genes, is required. However, filling the most prominent gaps in the global databases (e.g. GeneBank, PR2, BOLD), would be an acceptable alternative short-to-medium term solution.*

References

Aylagas E, Borja A, Rodríguez-Ezpeleta N (2014) Environmental status assessment using DNA metabarcoding: towards a genetic based marine biotic index (gAMBI). PLoS ONE 9(3): e90529.

Ji Y, Ashton L, Pedley SM, Edwards DP, Tang Y, Nakamura A, Kitchcing R, Dolman PM, Woodcock P, Edwards FA, et al. (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. Ecology Letters 16(10): 1245-1257

Evans NT, Olds BP, Renshaw MA, Turner CR, Li Y, Jerde CL, Mahon AR, Pfender ME, Lamberti GA, Lodge DM (2016) Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. Mol Ecol Resour. 16(1): 29-41.

Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, De Barba M, Gielly L, Lopes CM, Boyer F, et al. (2015) Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Molecular Ecology Resources 15: 543-556.

Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS ONE 6(4): e17497

Keeley N, Wood SA, Pochon X (2017). Development and validation of a multi-trophic metabarcoding biotic index for benthic organic enrichment biomonitoring. Ecological Indicators, (in review)

Lance KR, Carr MR 2012. Detecting eDNA of Invasive Dreissenid Mussels. Report on Capital Investment Project. ANSRP Technical Notes Collection. ERDC/TN ANSRP-12-2. Vicksburg, MS. U.S. Army Engineer research and Development Center, 9pp.

Laroche O, Wood SA, Tremblay LA, Ellis JI, Lear G, Pochon X 2017. A cross-taxa study using environmental DNA/RNA metabarcoding to measure biological impacts of offshore oil and gas drilling and production operations. Marine Pollution Bulletin (in press).

Lejzerowicz F, Esling P, Pillet L, Wilding TA, Black KD, Pawlowski J (2015) High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. Scientific Reports 5: 13932.

Pochon X, Cornelisen C, Morrisey D, Wood S 2016. Survey of microbial seabed communities in the Firth of Thames and Hauraki Gulf using a DNA metabarcoding approach. Prepared for Waikato Regional Council. Cawthron Report No. 2880. 22 p.plus appendix.

Pochon X, Wood SA, Keeley NB, Lejzerowicz F, Esling P, Drew J, Pawlowski J (2015) Accurate assessment of the impact of salmon farming on benthic sediment enrichment using foraminiferal metabarcoding. Marine Pollution Bulletin 100: 370-382

Ratnasingham S, Hebert PDN (2013) A DNA-based registry for all animal species: the barcode index number (BIN) system. PLoS ONE 8:e66213

Sboner A, Mu XJ, Greenbaum D, Auerbach RK, Gerstein MB (2011) The real cost of sequencing: higher than you think! Genome Biology 12: 125.

Shaw JLA, Clarke LJ, Wedderburn SD, Barnes TC, Weyrich LS, Cooper A (2016) Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. Biological Conservation 197: 131-138.

Wood SA, Smith KF, Banks JC, et al. (2013) Molecular genetic tools for environmental monitoring of New Zealand's aquatic habitats, past, present and the future. New Zealand Journal of Marine and Freshwater Research 47, 90-119

Yu DW, Ji YQ, Emerson BC, Wang XY, Ye CX, Yang CY, et al. (2012) Biodiversity Soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods Ecol. Avol. 3: 613-623.

Zaiko A, Samuiloviene S, Ardura A, Garcia-Vazquez E (2015) Metabarcoding approach for non-indigenous species surveillance in marine coastal waters. Marine Pollution Bulletin 100: 53-59.

Zaiko A, Pochon X, Garcia-Vazquez E, Olenin S, Wood S. Molecular tools for improved marine biosecurity: their applicability for deriving biodiversity information on pressures, ecosystem state and management success (in prep.).

4. Trait-based and functional diversity indicators. *Anna Törnroos (P2)*

Description

To better understand what may be gained and/or lost when we exploit the environment, we have turned the attention towards the goods (e.g. food and industrial revenues), and services (e.g. maintenance of biodiversity and favorable climate for humans) provided by the marine ecosystem (Holmlund & Hammer 1999, de Groot et al. 2002). In order to eventually assess these, we need to know the status and changes of the underlying ecosystem processes (e.g. bioturbation, bioirrigation, decomposition) and functions (e.g. primary production, nutrient cycling, carbon and organic matter pools), and inevitably how organisms and biological assemblages contribute to them.

Traditional taxonomical approaches have proven imperative for assessing both the status and changes in biological assemblages, however, they are limited in identifying the roles played by biological components and thereby any processes and/or functions of the ecosystem. Instead, embracing what organisms do as opposed to their taxonomy, is now widely accepted as the way to advance the field both in marine, freshwater and terrestrial realms. To address this, trait-based approaches (e.g. Biological Trait Analysis, single use of traits or functional groupings) have evolved and generated multiple different measures that conceptually have proven promising for further development into indicators (Beauchard et al. 2017). The inherent strength of trait-based approaches is their theoretical implications, that they allow for a mechanistic understanding of species performances in different environments (Southwood 1988).

A trait describes a particular characteristic of an organism, that can be either morphological, physiological, phenological or behavioral (Violle et al. 2007). Traits are preferably measured on an individual level, but practically and especially in marine environments more readily on a species level. Functional measures can essentially be divided into multi-trait ones, including several different traits, or single-trait ones, describing only one specific trait (Table 4.1). They can also be either weighted, by e.g. abundance or biomass, or be non-weighted, only present (expressed) or absent (not expressed). A trait is described either on a categorical or continuous scale. When a categorical scale is used, species are generally linked with traits through a fuzzy coding procedure (Chevenet et al. 1994).

Current measures essentially consider response traits, rather than effect traits (Dias and Cabido 1997, Hooper et al. 2004). Functional response traits are those traits which affect a species's response to changes in the environment, such as disturbance or resource availability, while effect traits are those which affect ecosystem properties (Lavorel and Garnier 2002). Response traits can vary independently from effect traits, and also indirectly affect functioning since the response to disturbance also affect sensitivity of organisms to recover from stress (Hooper et al. 2004). It is thus important to recognize this distinction and to some degree include also effect traits in any future indicator developments (Beauchard et al. 2017).

Another strength of trait-based indicators is that there is theoretically no restriction for which biodiversity component or ecosystem attribute it can address (Beauchard et al. 2017). However, most trait knowledge and assessments of trait-based measures are generally found for fish and zoobenthos (Beauchard et al. 2017), which is also reflected in the Baltic Sea (e.g. Törnroos et al. 2015, Gogina et al. 2014, Pecuchet et al. 2016, , Teixeira et al. 2016). There is, though, some work conducted on phytoplankton (Klais et al. 2016) and plants (Jänes et al. 2016), and even non-indigenous species (Aarnio et al. 2015) in the Baltic Sea. Trait-based measures also allow for cross-trophic analysis and can be developed to address food-web aspects, which has also been done in the Baltic Sea (Nordström et al. 2015).

Theoretically there are no restrictions to the geographic and/or temporal coverage of any future trait-based indicator. Practically though, the restrictions lie in the coverage of trait information (taxonomic level of assembled information, level of expertise of assigning traits) and the quantitative data and/or modelled possibilities of the biological component.

The functional aspect of ecosystems is mentioned in Article two of the Convention on Biological Diversity (CBD, 2001). Functional indicators also fall within the Marine Strategy Framework Directive (MSFD) purposes, particularly D1 and D6, but is also relevant for D4. As described above, no (multi-trait) functional measure has to date undergone a rigorous indicator assessment for the Baltic Sea that would target specifically any of the MSFD descriptors, but there exist some separate applications for single traits that are even identified as operational on a national level (Table 4, DevoTOOL). Measures under development are often based on the functional group concept or refer to “functional groups” although essentially only referring to one single trait (Table 4). However, most measures are only on a conceptual stage, still being scientifically evaluated (Table 4, DevoTOOL, Beauchard et al. 2017). No functional or trait-based indicators are currently included as a HELCOM CORE indicator, but scientific evaluations are under way in many working groups of ICES and/or HELCOM, e.g. ICES/HELCOM Working Group on Integrated Assessments of the Baltic Sea (WGIAB), that could potentially progress into suggestions of indicators in the future.

A number of natural and anthropogenic pressures have been related to trait-based measures globally, such as trawling (Tillin et al. 2006, de Juan and Demestre 2012, van Denderen et al. 2014), and gradients in pollution (Oug et al. 2012). In the Baltic Sea, trait-based measures have been assessed against pressures such as hypoxia (Gogina et al. 2014) and nutrient enrichment (Villnäs et al. 2011). Further assessments of pressures and their strength, and direction of change need to be undertaken.

Table 4.1. The hierarchy of trait-based measures with potential to be developed into indicators. Traits-based measures can be either weighted (with e.g. abundance or biomass) or non-weighted (presence/absence). They consist of either multiple traits or single traits. Important to remember is that a trait in any of these measures can be described either on a categorical or continuous level.

	Multi-traits	Single traits
Un-weighted	Functional groups (a number of different traits)*[1] Typological groups (a number of traits similar in performance)	E.g. Size*[2], diet*[3], life-form*[4] Bioturbation mode, Longevity*[5]

	Functional richness (based on n-dimensions in trait space), Trait richness (# of trait categories expressed)	
Weighted	Functional diversity, Functional evenness, Functional dispersion, Functional divergence	CWM (Community-weighted mean of a trait)*[5]

*[1]Abundance, composition and biomass of functional groups in i) selected habitats, and of ii) demersal invertebrates and fish are operational (according to DevoTOOL) in specific European areas (Armas et al. 2012, de Armas & Belles 2012, Khalaf 2002). An index for functional groups of benthic invertebrates is under development for the Baltic Sea (BSC 2008) as well as a seasonal succession of functional phytoplankton groups (HELCOM 2012). For references, please see DevoTOOL (search performed in October 2017, in version 6).

*[2]Body length distribution of selected benthic (long lived) species is under development in HELCOM and also as a Finnish national indicator (HELCOM 2012). For references, please see DevoTOOL (search performed in October 2017, in version 6).

*[3] Dietary functional group biomass of fish is under development in OSPAR, common indicator FW7 (ICES 2012, ICG-COBAM 2013). For reference, please see DevoTOOL (search performed in October 2017, in version 6).
*[4]Biomass ratio of plankton functional groups (based on life-forms) under development within OSPAR. For reference, please see DevoTOOL (search performed in October 2017, in version 6).

*[5] Utilizing the longevity distribution of a benthic community as a measure for trawling impact is under development in the North Sea (Rijnsdorp et al. 2016, Hiddink et al. 2017) and is being conceptually evaluated in the Baltic Sea (van Denderen et al. 2017 in prep.) as a preliminary proof- of concept. An indicator based on log number of long lived fish is also under development (Andersen et al. 2012)

*[5] Conceptual, could be used in e.g. temporal assessments (as extensions of Integrated trend analysis, ICES 2017).

Requirements for deriving the indicator
A number of trait-based measures exists, with ready-to-use, more or less standard analysis frameworks that could form a good basis for indicator development (Beauchard et al. 2017). However, currently used techniques comes with some caveats which needs further scrutiny, for example the dimensionality of analyzed patterns and fuzzy coding (Beauchard et al. 2017).

The minimum requirements for a fruitful development of trait-based indicators is trait information of the biological component and/or organism and preferably some quantitative information such as abundance and/or biomass. The latter is essentially provided through monitoring programs, but may need be tuned to assess a specific functional aspect. To date, trait information has primarily been collected on and used for adult stages, but larval and juvenile traits could also be included. In general, the Baltic Sea, with a relatively low number of common (often occurring) and dominant (in terms of abundance and/or biomass) species, present a potentially better case of trait collection than highly diverse systems. Although Baltic Sea-relevant trait information has been collected and to some degree published (e.g. Törnroos & Bonsdorff 2012).

A standardized trait database, spanning several taxonomic groups of Baltic Sea taxa does not exist to date. Such a database is critical and need to address both trait information (with references) and assignment of traits to species somehow, to ensure successful development and implementation of trait-based indicators.

Responsiveness to management measures
The responsiveness to management measures have not been assessed in the Baltic Sea and rarely outside either, and is thus a central aspect to tackle. Certain trait-based indicators respond differently than taxonomically-based ones along environmental gradients (Törnroos et al. 2015) and temporally (Törnroos et al. in prep.; Annex 3), and thus do not necessarily indicate a change when one is identified in the taxonomic structure. Whether this is of desire or not needs to be taken into consideration in indicator development as well.

Stability of the indicator
As described above, some initial explorative analysis exists of the spatio-temporal performance of some trait-based measures in the Baltic Sea, but greater and more targeted assessments related to pressures and management measures are essential to conduct.

Constrains and concerns
Care should be taken with regards to the selection of traits so that traits reflect the purpose of the potential indicator, which is essentially directed by the research or management question (Beauchard et al. 2017, ICES 2017). This is still a key issue although a number of measures show potential for future indicator development (Table 4.1). In reference to this, it is also suggested that specifically organism-based indicator development should aim to consider traits that both inform sensitivity and recoverability to pressures (Beauchard et al. 2017). Furthermore, while indicators derived from multiple traits are also suggested to provide better applicability as they reflect the theoretical aspects to a greater extent, they are the least applied in marine management (Beauchard et al. 2017).

Most studies assume a fixed value for a given species but traits generally vary considerably within species as well (Albert et al. 2011), and it is thus important to determine when this is not an issue for an indicator. Such information should also, preferably, be included in trait databases. In the marine realm, especially ontogeny but also spatial and temporal intra-specific trait variability should be considered.

Although the collection of trait information should essentially be a one-time investment, and thus cost-effective, some care should also need to be placed on trait adaptations to e.g. long-term pressures.

Advantages and Outlook
Trait-based indicators are believed to fill an essential gap in our understanding of changes in functional aspects of ecosystems to pressures and in the long run the provisioning of goods and services to us humans. With their wide applicability (biodiversity, functional and trophic aspects), trait-based measures thus have a great potential to better inform management of the Baltic Sea ecosystem.

As applications and assessments of trait-based measures are ever increasing, it is highly probable that technical constraints and some of the current constraints as well as empirical validations will be attended to, however more work is necessary for using current trait-based measures in decision-making tools.

With future refinement and validation of suggested trait-based indicator frameworks (e.g. Beauchard et al. 2017), delineating the pressure (non-human or human mediated), the state (effects on multiple traits) and responses (resistance, resilience or damage of traits), the building blocks for advancing Baltic Sea based ones are laid.

Summarized management advice

- *Trait-based measures have the advantage of addressing various functional aspects with greater mechanistic understanding and reference to our human needs and impacts on the marine ecosystem, compared to traditional taxonomical ones.*

- *The key steps in the development of trait-based measures into indicators of the Baltic Sea ecosystem should include:*

  (i) collection of trait data (novel or current existing data sets) into a trait data base,

  (ii) analysis of suitability of traits for different types of indicators and

  (iii) assessment of trait-based measures in indicator-based frameworks which delineates traits in relation to a specific pressure, the state of the system and responses to changes.

References

Aarnio K, Törnroos A, Björklund C, Bonsdorff E (2015) Food web positioning of a recent coloniser: the North American Harris mud crab *Rhithropanopeus harrisii* (Gould, 1841) in the northern Baltic Sea. Aquatic Invasions 10:399-413.

Beauchard O, Veríssimo H, Queirós AM, Herman PMJ (2017) The use of multiple biological traits in marine community ecology and its potential in ecological indicator development. Ecological indicators 76:81–96.

Chevenet F., Dolédec S, Chessel D (1994) A fuzzy coding approach for the analysis of long-term ecological data. Freshwater Biology 31:295-309.

de Groot RS, Wilson MA, Boumans RMJ (2002) A typology for the classification, description and valuation of ecosystem functions, goods and services. Ecological Economics 41, 393-408.

de Juan S, Demestre M (2012) A trawl disturbance indicator to quantify large-scale fishing impact on benthic ecosystems. Ecological Indicators 18:183-190.

Diaz S, Cabido M (1997). Plant functional types and ecosystem function in relation to global change. J. Veget. Sci. 8, 463-474.

ICES. 2017. Interim Report of the ICES/HELCOM Working Group on Integrated Assess- ments of the Baltic Sea. ICES WGIAB REPORT 2017 24-28 April 2017. Lisbon, Portugal. ICES CM 2017/SSGIEA:08. 32 pp.

Gogina M, Darr A, Zettler M (2014) Approach to assess consequences of hypoxia disturbance events for benthic ecosystem functioning. Jorunal of Marine Systems 129:203-213

Holmlund CM and Hammer M (1999) Ecosystem services generated by fish populations. Ecological Economics 29:253-268.

Hooper DU, Chapin FS, Ewel JJ, Hector A, Inchausti P, Lavorel S, Lawton JH, Lodge DM, Loreau M, Naeem S, Schmid B, Setala H, Symstad AJ, Vandermeer J, Wardle DA (2004) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecological Monographs 75:3-35

Jänes H, Kotta J, Pärnoja M, Crowe TP, Rindi F, Orav-Kotta H (2016) Functional traits of marine macrophytes predict primary production. Functional Ecology 31:975-986.

Klais R, Norros V, Lehtinen S, Tamminen T, Olli K (2016) Community assembly and drivers of phytoplankton functional structure. Functional Ecology 31:1-8.

Lavorel S, Garnier E (2002) Predicting changes in plant community composition and ecosys- tem functioning from plant traits: revisiting the Holy Grail. Functional Ecology 16, 545-556.

Nordström MC, Aarnio K, Törnroos A, Bonsdorff E (2015) Nestedness of trophic links and biological traits in a marine food web. Ecosphere 6:161

Oug E, Fleddum A, Rygg B, Olsgard F (2012) Biological traits analyses in the study of pollution gradients and ecological functioning of marine soft bottom species assemblages in a fjord ecosystem. J Exp Mar Biol Ecol 432:94–105

Pecuchet L, Törnroos A, Lindegren M (2016) Patterns and drivers of fish community assembly in a large marine ecosystem. Marine Ecology Progress Series, Elsevier, 546: 239-248

Rijnsdorp AD, Bastardie F, Bolam SG, Buhl-Mortensen L, Eigaard OR, Hamon KG, Hiddink JG, et al. (2016) Towards a framework for the quantitative assessment of trawling impact on the seabed and benthic ecosystem. ICES Journal of Marine Science: Journal du Conseil, 73: i127-i138.

Southwood TRE (1988) Tactics, strategies and templets. Oikos 52:3–18.

Teixeira H., Berg T., Uusitalo L., Fürhaupter K., Heiskanen A.-S., Mazik K., Lynam C.P., Neville S., Rodriguez J.G., Papadopoulou N., Moncheva S., Churilova T., Kryvenko O., Krause-Jensen D., Zaiko A., Verissimo H., Pantazi M., Carvalho S., Patricio J., Uyarra M.C., Borja A. 2016. A catalogue of marine biodiversity indicators. Frontiers in Marine Science 3: Article207.

Tillin HM, Hiddink JG, Jennings S, Kaiser MJ (2006) Chronic bottom trawling alters the functional composition of benthic invertebrate communities on a sea-basin scale. Marine Ecology Progress Series 318, 31–45.

Törnroos A., Bremner J., Bonsdorff E., Blomqvist M., Josefson A.B., Garcia C., Warzocha J. (2015) Marine benthic ecological functioning over decreasing taxonomic richness. Journal of Sea Research 98: 49-56.

Törnroos A, Bonsdorff E (2011) Developing the multitrait concept for functional diversity: Lessons from a system rich in functions but poor in species. Ecological Applications 22: 2221-2236.

Törnroos A, Olsson J, Gårdmark A, Pecuchet L, Lindegren M, Bonsdorff E (2017) Distinct functional dynamics, directional changes and interlinkages revealed in 40-year time series of marine fish and benthos. *In prep.*

van Denderen PD, Bolam SG, Hiddink JG, Jennings S, Kenny A, Rijnsdorp AD, and van Kooten T (2015). Similar effects of bottom trawling and natural disturbance on composition and function of benthic communities across habitats. Marine Ecology Progress Series, 541: 31-43.

van Denderen PD, Friedland R, Hiddink JG, Norén K, Rijnsdorp AD, Sköld M, Törnroos A, Valanko S (2017) Assessing impact of bottom trawl fishing and oxygen depletion on seafloor status. *In prep.*

Villnäs A, Perus J, Bonsdorff E (2007) Zoobenthos as indicators of ecological status in coastal brackish waters: a comparative study from the Baltic Sea. AMBIO 36:250-256

Violle C, Navas ML, Vile D, Kazakou E, Fortunel C, Hummel I, Garnier E (2007). Let the concept of trait be functional! Oikos 116:882–892.

5. **Benthic quality index.** *Anastasija Zaiko (P8), Joanna Calkiewicz, Jan Warzocha (P5)*

Description

Benthic Quality Index (BQI) is a widely used multimetric indicator of benthic community condition and functionality (Rosenberg et al. 2004, Fleischer et al. 2007, Fleischer and Zettler 2009, Leonardsson et al. 2009). BQI is reproducible and has been tested and validated in different marine ecosystems with varying environmental conditions (e.g. Labrune et al. 2006, Fleischer et al. 2007, Fleischer & Zettler 2009, Chuseve et al. 2016), therefore it was advised by the international expert groups (e.g. HELCOM CORESET) for distinguishing impacted habitats from undisturbed ones.

Although designed for application in marine areas (Borja et al. 2003, Rosenberg et al. 2004), it has proved to be suitable for areas with strong salinity gradients given that tolerance levels of species are properly adjusted and assigned for the specific area (Zettler et al. 2007). BQI has been already applied in a few Baltic Sea sub-regions.

The indicator is attributed to MSFD's Descriptor D6, but also relevant to D1, D4, D5, and D7 descriptors. It targets benthic invertebrates and is related to a number of pressures, including: physical loss, physical damage to marine habitats, interference with hydrological processes, contamination by synthetic compounds, contamination by non-synthetic substances and compounds, nutrient and organic matter enrichment, non-indigenous species and other biological disturbances.

There is no absolute and universal target value or GEnS boundary for BQI in the Baltic Sea, since natural species richness and distribution of benthic communities might significantly vary across ecosystems and different habitats.

BQI has been developed and considered operational in several Baltic sub-regions (Gulf of Finland, Bothnian Bay, The Quark, Bothnian Sea, Åland Sea, Northern Baltic Proper and Gotland Basin, Gulf of Riga, and Bay of Mecklenburg).

Requirements for deriving the indicator

The BQI approach has been developed through several consecutive studies (Rosenberg *et al*. 2004, Leonardsson *et al*. 2009, Leonardsson *et al*. 2015, Leonardsson *et al*. 2016, Blomqvist & Leonardsson 2016), and there are several versions of its calculation. Since the original version of BQI is known to be sampling effort dependent (e.g. increase in sampling effort results in higher probability of obtaining rare species), the adjusted calculation (Fleischer et al. 2007, Fleischer and Zettler 2009) is recommended:

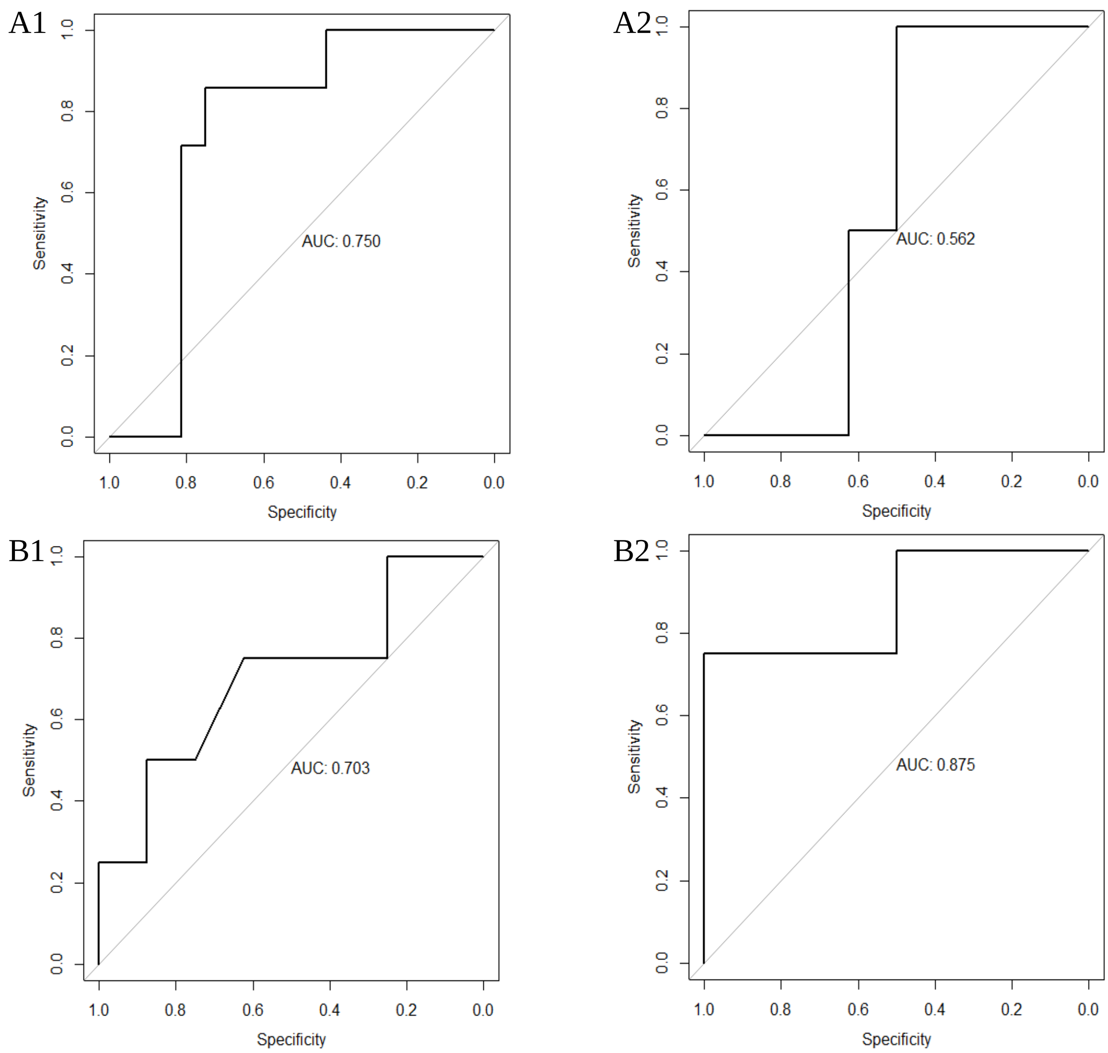
$$BQI_{ES} = \left(\sum_{i-1}^{n} \left(\frac{A_i}{A_{tot}} \times ES_{50_{0.05i}} \right) \right) \times \log(ES_{50} + 1) \times \left(1 - \frac{5}{5 + A_{tot}} \right)$$

In the equation above, *n* denotes the observed species number. A_i stands for the abundance of the species *i* and A_{tot} is the sum of all individuals within this square meter. Finally, $ES_{50-\,0.05}$ is the sensitivity/tolerance value for the species *i* and ES_{50} denotes the expected number of species for 50 individuals randomly taken from the square meter (Hurlbert Index).

For robust BQI assessment, the samples should be collected in a standardized manner, and an appropriate taxonomical expertise is required for consistent taxonomical assignment of detected organisms at the lowest possible level.

Responsiveness to management measures

The indicator does not directly linked to management response, since it might reflect multiple pressures/environmental factors present in the area. The suitability of the BQI indicator to represent response to the major pressures and relevant management measures, needed to be tested and validated using a standardized approach. For example, signal detection theory (SDT) was applied to evaluate the specificity and sensitivity of the BQI in SE Baltic Sea, its response to the eutrophication pressure and physical disturbance, and its performance under the effects of estuarine water outflow (Chuseve et al 2016). The BQI showed acceptable response to total nitrogen, total phosphorus and chlorophyll-*a* concentrations in the study area. The effect of the plume from the highly-eutrophicated Curonian Lagoon, have noticeably reduced sensitivity of the BQI to chlorophyll-*a* concentrations and specificity of response to the total phosphorus concentrations (Fig. 5.1).



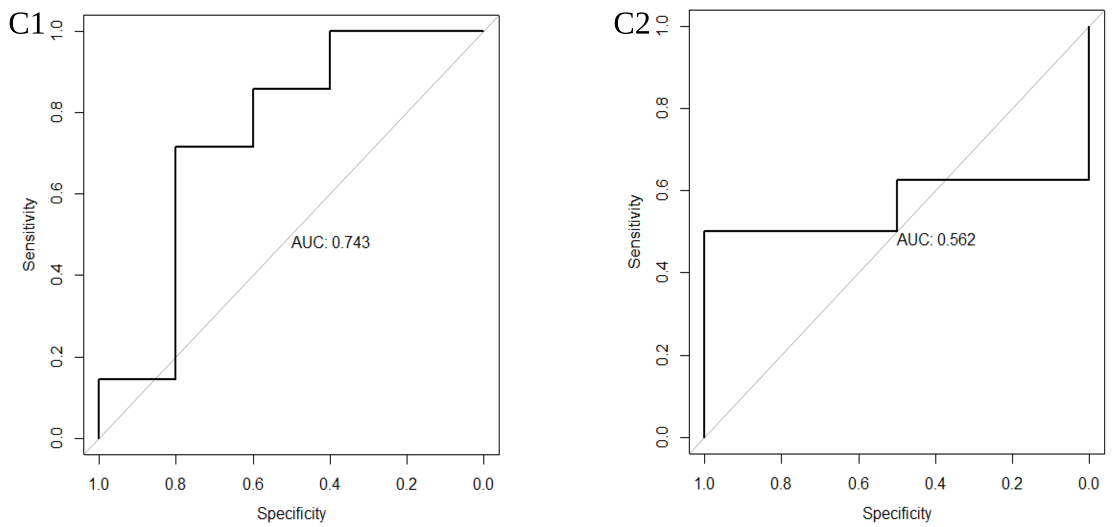


Figure 5.1. ROC (receiver operating characteristic) curves for BQI response to chl-a (A), TN (B) and TP (C) concentrations in the study area outside the plume zone (left column) and within the plume zone (right column).

The BQI sensitivity and response to untargeted pressures and "noise" disturbances was assessed on the dataset from the Curonian Lagoon (South-Eastern Baltic coast), highly affected by invasive zebra mussel *Dreissena polymorpha* (Zaiko & Daunys 2015). The results revealed that BQI values in samples with zebra mussels were significantly greater comparing to those devoid of zebra mussels, with no apparent temporal trend. There were some species recorded only in the presence of zebra mussel. Additionally, analysis of samples with zebra mussels demonstrated an evident effect of *D. polymorpha* abundance on the total macrofauna abundance, significantly correlating with a number of common soft-bottom species in the ecosystem. Therefore response of the BQI to the actual eutrophication-induced stressors (e.g. organic enrichments of sediment) was masked in habitats affected by *D. polymorpha* (Fig. 5.2). These findings were confirmed by the results from the southern Baltic (Polish EEZ). Although alien species did not change significantly the index values in the open waters, but the polychaete *Marenzelleria neglecta* strongly affected results in the Vistula Lagoon (Schiele et al. 2016).

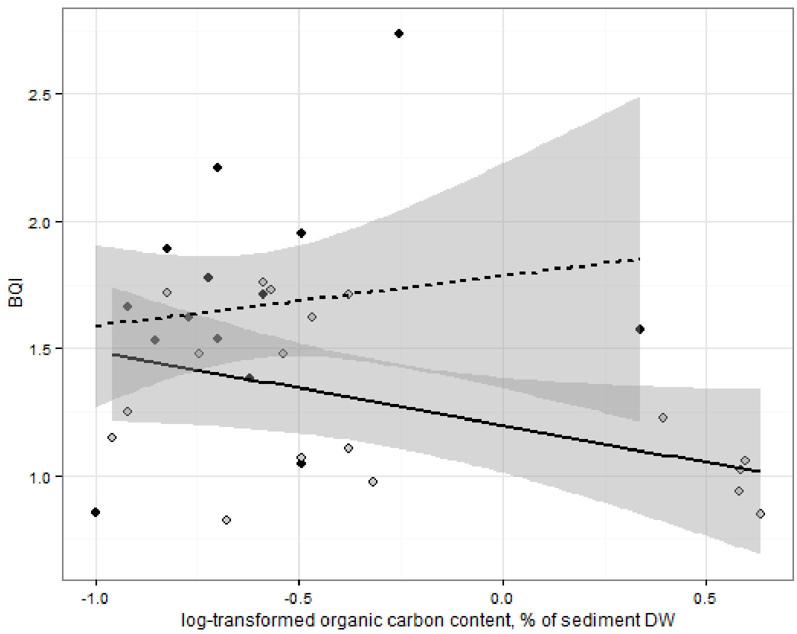


Figure 5.2: BQI values calculated for 1999 dataset based on the pre-assigned species sensitivity values. Grey dots – samples without zebra mussel: solid regression line (t=12; p<0.001; BQI= 1.19-

0.29x[log(C_org)]); black dots – samples with zebra mussel: dashed regression line ($t=3$; $p= 0.006$; BQI=1.89+0.36[log(C_org)]).

Based on the results of the Curonian Lagoon study, data correction framework was suggested to minimize the impact of the invasive species presence and abundance on BQI assessment (Zaiko & Daunys 2015).

Stability of the indicator

Since BQI targets benthic community, it can be considered rather temporarily stable, i.e. it is assumed that seasonal and short-term macrofauna variability of BQI should be rather low. However, there are no comprehensive empirical evidences on this from the Baltic sub-regions. The spatial stability can be affected by local habitat patchiness and peculiarities of benthic fauna distribution.

Thus, results of the indicator testing in the Polish EEZ suggest very considerable effect of natural habitat conditions, and particularly the variability of bottom sediments. The index best reflects the relatively homogenous muddy and sandy bottom in the Gulf of Gdańsk but, in the dynamic open sea area, it was affected by a patchy distribution of the sediment type. Additionally, irregular occurrence of mussel *Mytilus spp.* ,with associated small crustaceans, affected the indicator values as well. As a result it was significantly variable even within replicates at the same station. Extremely low index values were noted although habitats occurring there were probably the most similar to natural conditions. A different efficiency of the Van Veen grab on the muddy and sandy bottom also influenced the index value. This suggests the necessity to analyze the index in correspondence with habitat parameters. An appropriate location of sampling stations plays an important role as it allows the assessment of the anthropogenic pressure effects and will reflect the variability of local natural factors as much as possible. The use of sensitivity values calculated for different classes of salinity and depth (Schiele et al. 2016) allowed to reduce the masking effect of salinity and, to a certain extent, of different types of sediments.

More regional insight however is required on the small-scale spatial variability of BQI in response to different diffuse and point source pressures.

Constrains and concerns

Limited, variable and inconsistent taxonomic resolution can restrict ability of BQI to reliably represent benthic community response to the existing pressures and management measures.

As other benthic indices, BQI is based on the assumption that bottom-dwelling fauna are sedentary enough to escape from deteriorating environmental conditions and therefore will relatively rapidly respond to human induced pressures (Pearson and Rosenberg 1978; Borja et al. 2000). However, the integrative nature of its response to natural and anthropogenic pressures may restrict their ability to derive high signal to noise ratio and therefore effectively inform the management on appropriateness and effectiveness of the measurement measures applied.

Advantages and Outlook

BQI provides robust estimates of ecosystem health related to eutrophication and physical disturbance. Potentially, with the further development of the novel molecular techniques (such as eDNA metabarcoding), the BQI assessment could be refined and adjusted to such molecular data utilization, thus enhancing its sensitivity and accuracy.

Summarized management advice

- *Before routine exploitation of the BQI, a comprehensive analysis and fine-tuning of the calculation procedure should be undertaken for each particular geographical area/ecosystem/habitat. The assignment of sensitivities and target values should be also carefully revised and validated by the qualified scientists.*

- *Prior to routine application, the particular threshold values for BQI corresponding to GEnS should be set for the sub-region/habitat, based on a trial (preferably, long-term) dataset.*

- *The sensitivity values for the species should be also a-priory assigned to the major taxa, based on the trial dataset derived from the same habitat employing consistent and standardized sampling strategy.*

- *If there are invasive species in the ecosystem, affecting the benthic habitats and macrofauna communities, the BQI evaluation should be refined with appropriate corrections.*

- *To allow temporal-spatial comparisons, taxonomic resolution and naming should be aligned and standardized across sub-regions.*

- *Rapidly evolving molecular techniques (e.g. barcoding and metabarcoding) might be employed to complement and (eventually) substitute taxonomic identification process at least for some taxa. This would allow for better resolved and standardized taxonomic IDs*.

References

Blomqvist, M., Leonardsson, K. (2016). A probability based index for assessment of benthic invertebrates in the Baltic Sea. Deliverable D3.1-4, WATERS Report no. 2016:3 Havsmiljöinstitutet, Sweden.

Borja, A., Franco, J., Perez, V., (2000). A marine biotic index to establish the eco-logical quality of soft-bottom benthos within European estuarine and coastalenvironments. Mar. Pollut. Bull. 40, 1100–1114.

Borja A, Muxika I, Franco J (2003) The application of a Marine Biotic Index to differ-ent impact source affecting soft-bottom benthic communities along Europeancoasts. Mar. Pollut. Bull. 46, 835–845.

Chuseve R, Nygard H, Vaiciute D, Daunys D, Zaiko A (2016) Application of Signal Detection Theory approach for setting thresholds in benthic quality assessments. Ecological Indicators, 60: 420-427.

Fleischer D, Gremare A, Labrune C, Rumohr H, Berghe EV, Zettler ML (2007) Performance comparison of two biotic indices measuring the ecological statusof water bodies in the Southern Baltic and Gulf of Lions. Mar. Pollut. Bull. 54, 1598–1606.

Fleischer D, Zettler ML (2009) An adjustment of benthic ecological quality assess-ment to effects of salinity. Mar. Pollut. Bull. 58, 351–357.

Labrune C, Amouroux JM, Sarda R, Dutrieux E, Thorin S, Rosenberg R, Gremare A (2006) Characterization of the ecological quality of the coastal Gulf of Lions(NW Mediterranean). A comparative approach based on three biotic indices.Mar. Pollut. Bull. 52, 34–47.

Leonardsson K, Blomqvist M, Rosenberg R (2009) Theoretical and practical aspectson benthic quality assessment according to the EU-Water Framework Directive– examples from Swedish waters. Mar. Pollut. Bull. 58, 1286–1296.

Leonardsson, K., Blomqvist, M., Magnusson, M., Wikström, A., Rosenberg, R. (2015). Calculation of species sensitivity values and their precision in marine benthic faunal quality indices. Mar. Pollut. Bull. 93, 94-102.

Leonardsson, K., Blomqvist, M., Rosenberg, R. (2016). Reducing spatial variation in environmental assessment of marine benthic fauna. Mar. Pollut. Bull. Doi: 10.1016/j.marpolbul.2016.01.050.

Pearson, T.H., Rosenberg, R., (1978). Macrobenthic succession in relation to organicenrichment and pollution of the marine environment. Oceanogr. Mar. Biol. Annu.Rev. 16, 229–311.

Rosenberg R, Blomquist M, Nilsson HC, Cederwall H, Dimming A (2004) Marinequality assessment by use of benthic species-abundance distributions: a pro-posed new protocol within the European Union Water Framework Directive.Mar. Pollut. Bull. 49, 12.

Schiele K. S., Darr A., Zettler M., Berg T., Blomqvist M., Daunys D,, Yermakovs V., Korpinen S., Kotta J., Nygård H., von Weber M., Voss J.,Warzocha J. 2016. Rating species sensitivity throughout gradient systems – a consistent approach for the Baltic Sea. Ecological Indicators. 61, 447-455.

Zaiko A, Daunys D (2015) Invasive ecosystem engineers and biotic indices: giving a wrong impression of water quality improvement? Ecological Indicators 52: 292-299.

Zettler ML, Schiedek D, Bobertz B (2007) Benthic biodiversity indices versussalinity gradient in the southern Baltic Sea. Mar. Pollut. Bull. 55, 258–270.

6. Zooplankton Mean Size and Total Stock (MSTS) indicator. *Piotr Margonski and Joanna Calkiewicz (P5), Maiju Lehtiniemi (P7)*

Description

This core HELCOM indicator is mainly relevant for food webs (EU Marine Strategy Framework Directive (MSFD) criterion 4.3: *abundance/distribution of key trophic groups/species*) with a secondary link to biodiversity (EU MSFD criterion 1.6: *habitat condition*) but according to the newly-adopted Commission Decision (EU) 2017/848) on laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardised methods for monitoring and assessment, MSTS became relevant primarily to criterion **D4C3** (*the size distribution of individuals across the trophic guild is not adversely affected due to anthropogenic pressures*) with a secondary link to **D1C6** (*the condition of the habitat type, including its biotic and abiotic structure and its functions is not adversely affected due to anthropogenic pressures*).

The zooplankton mean size and total stock (MSTS) trends indicate that the investigated pelagic food web structure is or is not optimal for energy transfer from primary producers (phytoplankton) to fish and for regulating the phytoplankton community (Gorokhova et al. 2016).

Pivotal position of zooplankton organisms in the pelagic food web is indicative of both fish feeding conditions (and to some extent the predatory pressure of small pelagic fish on zooplankton) as well as grazing pressure of zooplankton on phytoplankton (HELCOM 2015). MSTS is strongly linked to two anthropogenic pressures listed in the MSFD Annex III, Table 2: selective extraction of species and nutrient and organic matter enrichment.

As described in HELCOM (2015), this core indicator employs zooplankton mean size (MS) and total stock (TS) to evaluate pelagic food web structure, with particular focus on lower levels. MSTS evaluates good environmental status (GES) using two GES boundaries, one for mean size and one for total standing stock (abundance or biomass) of zooplankton. The mean zooplankter size is presented as a ratio between the total zooplankton abundance (TZA) and total biomass (TZB). This metric is complemented with an absolute measure of total zooplankton stock, TZA or TZB, to provide MSTS. Thus, MSTS is a two-dimensional or multimetric indicator representing a synthetic descriptor of zooplankton community structure (Figure 6.1).

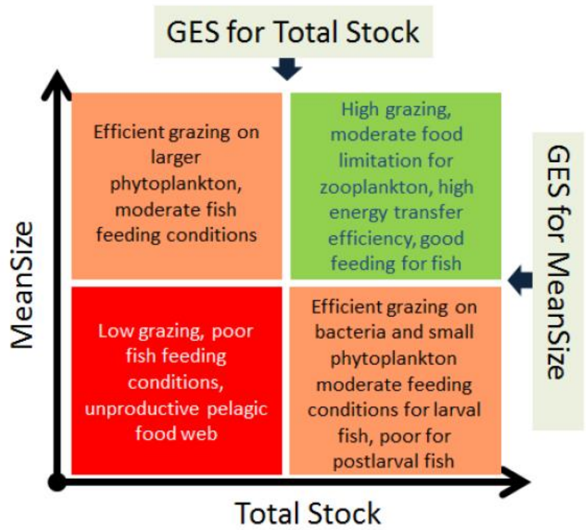


Figure 6.1. The MSTS concept. The green area represents GES condition (i.e. threshold values), yellow areas represent sub-GES conditions where only one of the two parameters is adequate and the red area represents sub-GES conditions where both parameters failed (HELCOM 2015).

HELCOM (2015) states that the GES boundaries (i.e. threshold values) are set using a reference period within existing time series that defines a status when the food web structure was not measurably affected by eutrophication and represents good fish feeding conditions. This indicator evaluates the structural and functional integrity of the food web. Thus, on a conceptual level GES (i.e. threshold value) is achieved when:
there is a high proportion of large-sized individuals (mostly copepods) in the zooplankton community that efficiently graze on phytoplankton and provide good-quality food for zooplanktivorous fish, and
the abundance (biomass) of zooplankton is at an adequate level to support fish growth and exert control over phytoplankton production.

The reference periods for MSTS should reflect a time period when effects of eutrophication, defined as 'acceptable' chlorophyll *a* concentration, are low, whereas nutrition of zooplanktivorous fish is adequate for optimal growth. Hence, these are the periods when eutrophication- and overfishing-related food web changes are negligible. Target setting is based on actual data from the reference period within existing monitoring data series for the respective areas.

Requirements for deriving the indicator
A detailed description of methods to calculate the MSTS threshold values as well as the assessment protocol has been provided by HELCOM (2015) report. The indicator evaluation is currently limited to the summer time period (June-September) which is considering the period of the highest plankton productivity, high predations pressure on zooplankton, as well as the current monitoring data availability.

The two approaches to setting the reference period are suggested by HELCOM (2015) report: (1) all data available (i.e. all years of the monitoring period), and (2) a window of the available data corresponding to the selected reference period (i.e., years representing basin-specific reference conditions for (i) food webs not measurably affected by eutrophication; these are based on environmental quality ratio (EQR) and historical data on chlorophyll *a* (HELCOM 2009) when defining $RefCon_{Chl}$, and (ii) high feeding conditions for zooplanktivorous fish when defining $RefCon_{Fish}$. The mean and standard deviation of the z-score is defined based on the reference period conditions.

When testing the MSTS concept (Gorokhova et al. 2016) within the BIO-C3 project and developing the reference periods based on the Polish monitoring data from the southern Baltic Sea, we defined the relevant reference periods using modeled data: (i) a chlorophyll *a* EQR was calculated from the RCO-

SCOBI model (Meier et al., 2012) output for the zooplankton sampling station location and (ii) $RefCon_{Fish}$ were calculated based on sprat average biomass in age groups from the Stochastic Multi-Species model output (Lewy and Vinther, 2004) for ICES subdivisions 25 or 26.

Responsiveness to management measures

MSTS is strongly linked to two anthropogenic pressures listed in the MSFD Annex III, Table 2: selective extraction of species and nutrient and organic matter enrichment. Therefore, it is designed to show a quick and positive response to any successful management measures affecting the actual level of those two pressures.

Stability of the indicator

The MSTS indicator is stable in terms of threshold values when it is calculated based on properly identified reference conditions. The other option exists according to the indicator description (HELCOM, 2015) when available monitoring data set is shorter and then all data available (i.e. all years of the monitoring period) are used for setting the reference period. In this case, adding new data in the following year may influence the calculated threshold value.

Spatial stability will be dependent on the local dynamics in zooplankton community between the sub-regions that are compared.

Constrains and concerns

The MSTS indicator is useful but only when calculated based on data series long enough to at least partly represent the reference conditions. Only in such a case the calculated threshold values will be stable and representative of the GES conditions. Moreover, zooplankton biomass estimates have to be unified over the entire Baltic Sea region (this work has already started in coordination of the HELCOM Zooplankton Expert Network (ZEN ZIIM) group).

Advantages and Outlook

The MSTS indicator appears to be a promising tool to test and evaluate the conditions and the spatio-temporal dynamics of the pelagic food web considering both bottom-up as well as top-down impacts at lower food web levels.

Summarized management advice

- *Biomass estimates of zooplankton individuals have to be unified over the entire Baltic Sea.*
- *Reference conditions for chlorophyll a EQR and sprat and herring condition should be identified based on the sub-regional (i.e. regionally disaggregated) data to reflect the spatial dynamics at the local scale.*
- *Before setting the final threshold values the inter-calibration exercise has to be conducted in all the regions or sub-regions where more than one monitoring data series exists.*

References

Gorokhova, E., M. Lehtiniemi, L. Postel, G. Rubene, C. Amid, J. Lesutiene, L. Uusitalo, S. Strake, and N. Demereckiene. 2016. Indicator Properties of Baltic Zooplankton for Classification of Environmental Status within Marine Strategy Framework Directive. PLoS ONE **11**:e0158326

HELCOM 2009. Eutrophication in the Baltic Sea – An integrated thematic assessment of the effects of nutrient enrichment and eutrophication in the Baltic Sea region. Balt. Sea Environ. Proc. No. 115B

HELCOM 2015. Zooplankton Mean Size and Total Stock (MSTS). HELCOM core indicator report. Online. http://www.helcom.fi/Core%20Indicators/Zooplankon%20mean%20size%20and%20total%20stock-HELCOM%20core%20indicator%20report%202015-extended%20version.pdf. Document viewed on 2017-10-12.

Lewy, P., and Vinther, M. 2004. Modelling stochastic age-length-structured multi-species stock dynamics. ICES C.M. 2004/FF:20: 1-33.

Meier, H. E. M., Andersson, H. C., Arheimer, B., Blenckner, T., Chubarenko, B., Donnelly, C., Eilola, K., et al. 2012. Comparing reconstructed past variations and future projections of the Baltic Sea ecosystem-first results from multi-model ensemble simulations. Environmental Research Letters, 7(3): 034005.This section holds the references for the citations used in the text of the individual indicator.

7. Food web and phytoplankton indicators. *Harri Kuosa (P7), Monika Winder (P4), Riina Klais (P6)*

Description

This indicator is under development. It has been introduced as one of the 'candidate indicators' in HELCOM. A full description is given in Lehtinen et al. (2016).

The indicator is derived from phytoplankton monitoring data to draw conclusions about the potential effect of phytoplankton taxonomic composition on the next trophic level. This information can be used as a part of marine food web assessments required by the Marine Strategy Framework Directive of the European Union. In the indicator development, both contemporary taxonomic composition and recent trends are used to assess the potential consequences for food web functioning. The calculation of the indicator consists of four steps: (1) long-term trend analysis of class-level and total phytoplankton biomass using generalized additive models (GAMs) and calculation of mean biomass proportion of each phytoplankton class from the total phytoplankton biomass, (2) comparison of the current phytoplankton community composition against its long-term variability with non-metric ordination analysis (NMDS), applied to the genus-level biomass, (3) identification of the dominant taxa (to the lowest taxonomic level), as well as the taxa that drive the long-term variability in the community composition, and (4) interpretation of the results from 1-3 to suggest the potential effects on the next trophic level. In step 4, special attention is given to the dominant and most time-variable taxa, in particular to the following ecological traits: suitability and quality as food for grazers, harmfulness, size, and trophy. These characteristics are selected based on existing scientific knowledge on their relevance to the higher trophic levels.

A healthy phytoplankton community supports the abundant and diverse micro- and mesozooplankton communities, which in turn for the basis of healthy fish communities. Phytoplankton is the first trophic level directly and instantly responding to the changes in nutrient availability. In addition to external nutrient loading, phytoplankton composition responds to internal nutrient loading, physical conditions, and food web interactions.

The monitoring of phytoplankton community composition is essential, since the composite indicators, e.g. chlorophyll-a (chl-a) and the occurrence of algal surface accumulations, are not sufficient to describe quality of phytoplankton community from the food web perspective. This is mainly because the chl-a measurements and surface accumulation observations do not provide information on the species composition, e.g. whether the phytoplankton community is dominated by toxic or nontoxic species, or by high or low-quality food. Also the primary production estimates derived via remote sensing have been suggested to be used, especially as an estimate of the resources available for the fisheries (ICES 2014), but this is even more far fetched, considering that even chl-a concentration measured from water samples can not predict the effect of primary producers to the food web structure and healthiness, or on zooplankton and fish (Suikkanen et al. 2013, Kuosa et al. 2017). For example, phytoplankton community composition during the spring bloom is a very important factor affecting the bio-geochemical cycles of the Baltic Sea (Spilling & Lindström 2008, Klais et al. 2011). Thus, phytoplankton community composition should be used to assess the quality of the primary production, not only quantity.

Phytoplankton fatty acid composition varies according to phytoplankton taxonomical groups, therefore the quality of phytoplankton as a food source in aquatic habitats is primarily based on the dominating groups (Galloway and Winder 2015). The production of high-quality food is shown to support higher secondary production in coastal ecosystems (Winder et al. 2017).

Requirements for deriving the indicator

Robust harmonized methods for sampling, microscopy, and biomass calculations are required for deriving the indicator.

In the test, sampling frequency was once a year during late July to early September. The methodology follows the HELCOM COMBINE Manual (HELCOM 2015): Samples were taken as integrated water samples from the surface layer (0-10 m) by mixing equal amounts of water from the depths 1, 2.5, 5, 7.5, and 10 m. Samples were preserved with acidic Lugol's solution (1 ml Lugol's per 300 ml sample), and kept refrigerated (+4 - +10ºC) and in the dark before the microscopic analysis within a year of sampling. Microscopy was performed with an inverted light microscope using the Utermöhl method. A volume of 50 ml (25 ml) of sample was settled into the settling chamber. A magnification of 125x was used to count the >30 µm –sized and sparse species, as well as Nostocales. A magnification of 250x was used to count the 20-30µm –sized species, Chroococcales colonies with a cell size of >2 µm, as well as Oscillatoriales. A magnification of 500x was used to count <20 µm –sized species and Chroococcales colonies with a cell size of <2µm. 60 ocular squares were analyzed with each of the three magnifications, aiming to count at least 400 counting units with each magnification. The counting units and size classes of the HELCOM PEG taxa and biovolume list were used to convert the results into biomasses µg/l (Olenina et al 2006, a link to the annually updated Biovolume file is available on the HELCOM PEG www-page http://helcom.fi/helcom-at-work/projects/phytoplankton).

Taxonomic nomenclature and biomass results should be harmonized, e.g. with the latest version of the HELCOM PEG taxa and biovolume list. A link to the annually updated Biovolume file is available at http://helcom.fi/helcom-at-work/projects/phytoplankton.

Responsiveness to management measures

Suikkanen et al. (2013) concluded that the observed changes in plankton communities are most likely due to the interactive effects of climate warming, eutrophication and increased top-down pressure due to overexploitation of resources, leading to the trophic cascades.

Stability of the indicator

Due to the high spatiotemporal variability of phytoplankton, low sampling frequency can be the decisive caveat for deriving robust indicator assessment. A higher sampling frequency during the season would increase the accuracy, and shorten the monitoring period which is necessary to be able to detect the significant temporal changes. Thus, the commendatory sampling frequency should be at least monthly, but preferably even twice a month, and covering the whole year.

Constrains and concerns

The implementation of the indicator is currently impeded by the lack of full description of phytoplankton functional characteristics, the collection of functional traits to complete the species list is ongoing. Also, the reference conditions are difficult to subtract from the existing time series. Threshold value calculation requires high-quality long-term data from different sea areas.

Advantages and Outlook

The indicator based on phytoplankton community composition is a promising tool to test and evaluate the bottom-up effects of primary producers to the higher levels of food web. It also allows the inclusion of field and experimental research on phytoplankton to a common frame, which widens the view of primary producers from a black box towards a more detailed community with a multitude of functions.

Summarized management advice

- *The sampling of the existing long-term monitoring stations should be continued with the existing robust methods, but the sampling frequency should be increased preferably to twice a month during the whole year.*

- *In addition to phytoplankton community composition, other trophic levels should be monitored as well (sampling frequency varies for different parameters) to be able to perform a holistic assessment of food web functioning and healthiness, and to follow the synchronous changes across the trophic levels.*

References

Galloway AWE, Winder M 2015. Partitioning the Relative Importance of Phylogeny and Environmental Conditions on Phytoplankton Fatty Acids. PLoS ONE 10(6): e0130053. doi:10.1371/journal.pone.0130053

Kuosa H, Fleming-Lehtinen V, Lehtinen S, Lehtiniemi M, Nygård H, Raateoja M, Raitaniemi J, Tuimala J, Uusitalo L, Suikkanen S 2017. A retrospective view of the development of the Gulf of Bothnia ecosystem. Journal of Marine Systems 167 (2017) 78–92

Lehtinen S, Suikkanen S, Hällfors H, Kauppila P, Lehtiniemi M, Tuimala J, Uusitalo L, Kuosa H 2016. Approach for supporting food web assessments with multi-decadal phytoplankton community analyses – case Baltic Sea. Front. Mar.Sci.3:220. doi: 10.3389/fmars.2016.00220

HELCOM, 2015. Guidelines concerning phytoplankton species composition, abundance and biomass. Manual for Marine Monitoring in the COMBINE Programme of HELCOM. Annex C-6, Last updated February 2015. http://www.helcom.fi/action-areas/monitoring-and-assessment/manuals-and-guidelines/combine-manual

Klais R, Tamminen T, Kremp A, Spilling K, Olli K 2011. Decadal-Scale Changes of Dinoflagellates and Diatoms in the Anomalous Baltic Sea Spring Bloom. PLoS ONE 6(6): e21567.

Rogers S, Casini M, Cury P, Heath M, Irigoien X, Kuosa H, Scheidat M, Skov H, Stergiou K, Trenkel V, Wikner J, Yunev O 2010. Marine Strategy Framework Directive – Task Group 4 Report Food webs. Scientific and Technical Research series. EUR 24343 EN – Joint Research Centre. European Union and ICES. Luxembourg: Office for Official Publications of the European Communities. 55 pp.

Spilling K, Lindström M. 2008. Phytoplankton life cycle transformations lead to species-specific effects on sediment processes in the Baltic Sea. – Continental Shelf Research 28(17):2488–2495.

Suikkanen, S., Pulina, S., Engström-Öst, J., Lehtiniemi, M., Lehtinen, S., Brutemark, A. 2013. Climate change and eutrophication induced shifts in Northern Summer plankton communities. PLOS ONE 8(6) e66475.

Winder M, Carstensen J, Galloway AWE, Jakobsen HH, Cloern JE 2017. The land–sea interface: A source of high-quality phytoplankton to support secondary production. Limnology and Oceanography doi: 10.1002/lno.10650

8. Genetic indicators. *Thorsten Reusch (P1)*

Description

Genetic indicators broadly address two different goals. Indicators based on neutral genetic diversity are suited to track population identity and population mixing, important for example, in tracking invasion routes (Reusch et al. 2010; Bio-C3 Deliverable D1.3) or identifying distinct fish stocks (Nielsen et al. 2012). A second suite of genetic indicators addresses adaptive genetic diversity, and thus, would allow predictions if the genetic diversity that is prerequisite for adaptive evolution under global change is available (hereafter selection-based genetic indicators; see e.g. Nielsen et al. 2009)

For selected fish species and also invertebrates, neutral genetic indicators are established, for example through the EU project FishPopTrace (Nielsen et al. 2013), or via other research projects for dedicated invasive species (Reusch et al. 2010).

Deriving meaningful indicators for adaptive genetic diversity is much more challenging because several intermediate steps need to be resolved. First, the link between genotype and phenotype of a particular relevant trait would need to be established. Second, a robust genotyping systems needs to be established and verified. Work in this direction is ongoing, in particular in selected fish species where full genome information plus several re-sequenced genomes are available, permitting to resolve the genotype-phenotype correlation (e.g. Hemmer-Hansen et al. 2013).

As a shortcut, neutral genetic diversity can be taken as a surrogate for the general diversity at selectively relevant sites, if the results are interpreted with caution. For example, it has been shown that heavily exploited populations of commercially important fish show less genetic variation at microsatellite and SNP sites (Pinsky and Palumbi 2014).

Requirements for deriving the indicator

Both major types of genetic indicators, both neutral and selection based, need species-specific and sometimes even population-specific development that requires an initial collection of multiple genomes or genome fragments of the target species /population. In addition, the selection-based markers also need time-consuming development to establish a link to the putative functionality of the identified genetic diversity.

In contrast, the technical side of marker establishment, once the functional relationships have been addressed, is becoming more and more straightforward owing to the rapid progress of sequencing and genotyping technologies. A clear recommendation for any future genetic indicator entails its full reproducibility across laboratories, which will be greatly enhanced by using SNP (single nucleotide polymorphism) genotyping technology that provides us, analogous to DNA sequencing, with digital information, in contrast to microsatellite length polymorphism.

Responsiveness to management measures

As for neutrality-based markers that track population identity, there is a wealth of information from fish populations that proved indispensible for fisheries management in the Baltic Sea region. Within BIO-C3, it was found that east and west stocks of Baltic cod meet as adults in the Arkona basins, but that the do not mix (Hemmer-Hanssen et al. unpubished; Bio-C3 Deliverable D1.3), and that the division between both stocks is maintained without hybridization for at least 2 decades (Dierking et al. unpublished, Bio-C3 Deliverable D1.3)

Stability of the indicator

Neutrality based genetic indicators should be stable to track management units over years to decades. Selection-based genetic indicators will need recalibration under changing environment, since the allele frequencies change under adaptive evolution, they are thus a "moving" target.

Constrains and concerns
While neutrality based genetic indicators are undisputed, it is as yet unclear what the optimal strategy would be to address adaptive relevant genetic diversity using selection-based genetic indicators.

Diversity information is always a "byproduct" of any neutrality-based assessment. It is hence certainly a good idea to compute basic measures of genetic diversity at neutral markers, such as heterozygosity and the effective number of alleles, and compute their time course. Under the precautionary approach, any reduction in neutral genetic diversity should be a matter of deep concern (Spielmann et al. 2004), as this may endanger the capacity for populations to adapt to a changing environment (Reusch 2014).

Advantages and Outlook
Genetically based markers of the "new" generation, i.e. those that are DNA-sequence based, are digital and hence, completely inter-calibrated over longer time scales of observation. Their acquisition will become cheaper and cheaper. The bottleneck at the moment is assigning function to their polymorphism. We envisage that functional knowledge will accumulate over the coming decades in many species and their important trails, so that it will become feasible to monitor the availability of adaptive variation via genetic indicators directly. As a recommendation, sampled tissue should be preserved in dedicated bio-banks for long-term genetic monitoring even if sensible functional indicators are currently not yet in place.

Summarized management advice

- *Genetic diversity indicators are promising for:*
  - *tracking population identity and population mixing (e.g. to identify routes of bioinvasions thus improving certainty regarding introduction pathways, or to identify distinct fish stocks)*
  - *assessing adaptive genetic diversity (e.g. to predict if the adaptive evolution under global change)*
- *These indicators need dedicated effort for development and establishing baselines*
- *Working forward to improve knowledge on functional traits relatedness to genetic polymorphism*
- *Systematic collection of genetic information (with appropriate long-term sample storage for reference) would allow to monitor the availability of adaptive variation via genetic indicators directly, as the functional knowledge improves*

References
Hemmer-Hansen J, Nielsen EE, Therkildsen NO, Taylor MI, Ogden R, Geffen AJ, Bekkevold D, Helyar S, Pampoulie C, Johansen T, FishPopTrace C, Carvalho GR (2013) A genomic island linked to ecotype divergence in Atlantic cod. Molecular Ecology 22:2653-2667

Nielsen EE, Cariani A, Aoidh EM, Maes GE, Milano I, Ogden R, Taylor M, Hemmer-Hansen J, Babbucci M, Bargelloni L, Bekkevold D, Diopere E, Grenfell L, Helyar S, Limborg MT, Martinsohn JT, McEwing R, Panitz F, Patarnello T, Tinti F, Van Houdt JKJ, Volckaert FAM, Waples RS, FishPopTrace c, Carvalho GR (2012) Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. Nat Commun 3:851

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009) Population genomics of marine fishes: identifying adaptive variation in space and time. Molecular Ecology 18:3128-3150

Pinsky ML, Palumbi SR (2014) Meta-analysis reveals lower genetic diversity in overfished populations. Mol Ecol 23:29-39

Reusch, T., Dierking, J., Hemmer-Hansen, J., Karlsson, Konrad.... (in review) *Report on adaptive evolution linking trait and functional genetic variance for selected species*. BIO-C3 Deliverable, D1.3. EU BONUS BIO-C3, DOI xxxx

Reusch TBH (2014) Climate change in the oceans: evolutionary vs. phenotypically plastic responses in marine animals and plants. Evolutionary Applications 7:104–122

Reusch TBH, Bolte S, Sparwel M, Moss A, Javidpour J (2010) Microsatellites reveal origin and genetic diversity of Eurasian invasions by one of the world's most notorious marine invader, *Mnemiopsis leidyi* (Ctenophora). Mol Ecol 19:2690–2699

Spielmann D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences USA 101:15261-15264

Appendices

Annex 1: Report on response of biodiversity indicators to management measures (test of indicators). ***Andrea Rau (P11)***

Annex 2: Manuscript on molecular tools for improved marine biosecurity - ***not for public disclosure until published.*** Corresponding author: Anastasija Zaiko (P8) anastasija.zaiko@jmtc.ku.lt

Annex 3: Manuscript on Long-term functional change in benthos and fish - ***not for public disclosure until published*****.** Corresponding author: Anna Törnroos (P2) anna.m.tornroos@abo.fi

Annex 1. Report on response of biodiversity indicators to management measures (test of indicators).

P11: Test of MSFD indicators concerning (commercial) fish species

Authors: Andrea Rau[1],
[1] Thünen-Institute of Baltic Sea Fisheries

1 Introduction

In 2008 the European Marine Strategy Framework Directive (MSFD) came into force (EU-COM 2008), aiming for good environmental status (GES) in all European marine waters by 2020. In the following for the ongoing assessment of marine environmental status the European Commission provided member states with a document (2010/477/EU) containing a list of environmental indicators (EU-COM 2010). The here presented study tested the performance of those MSFD indicators related to fish and was conducted by the Thünen-Institutes of Sea Fishery and Baltic Sea Fisheries in Germany (Probst et al. 2014). The focus was on the German use of indicators for its national marine waters in the North Sea and Baltic Sea. The indicators were evaluated by using a scoring system enabling objective and quantitative expert estimation on the quality of the indicators against certain specified criteria following consistent guidelines.

1.1 List of selected indicators

Table 1 shows the indicators selected for testing, all of them are related to fish and refer to the commission decision document 2010/477/EU of 1st September 2010 on criteria and methodological standards on good environmental status of marine waters. Commercial fish are explicitly treated under Descriptor 3 (D3) while more generally fish are included also within Descriptor 1 (D1) concerning biodiversity aspects and within Descriptor 4 (D4) relating to fish community aspects. According to ICES advice on draft recommendations for the assessment of MSFD Descriptor 3 from March 2014 (ICES 2014) the EU-COM D3 indicators 3.3.2 (mean maximum length across all species) and 3.3.4 (size at first sexual maturation) were not considered for evaluation of indicator performance. D1 indicators on distributional range (1.1.1) and distributional pattern (1.1.2) are combined according to the German approach. In addition, two indicators are analyzed which are not part of the EU-COM decision document on criteria and standards – the indicator 1.3.1 "Bycatch/discard of selected fish species (target and non-target species, e.g. vulnerable species) in relation to population size" and the indicator 1.4.1 "Conservation status of fish" are proposals on the German list of indicators only.

2 Method of testing

2.1 Assessment of indicator performance

An indicator scoring was conducted based on expert knowledge to evaluate the fish related MSFD indicators, following the evaluation methodology of the ICES Working Groups on Biodiversity Science (WGBIODIV) and on Ecosystem Effects of Fishing Activities (WGECO) who specifically developed a set of criteria to assess performance of indicators to support implementation of the MSFD (ICES 2013). The indicators were rated against 15 criteria referring to quality of underlying data (e.g. data availability, sensitivity to pressures), importance for management (e.g. relevance to MSFD objectives, comprehensibility, cost effectiveness) and conceptual performance (e.g. theoretically soundness, relevance to MSFD indicators). Accordingly indicator performance was evaluated considering three aspects: (I) Overall evaluation, taking into account all criteria; (II) evaluation of operationality,

Table 1: List of selected MSFD indicators referring to fish

Descriptor	Criterion	Indicator
D1: Biodiversity	1.1 Species distribution	1.1.1/1.1.2 Distributional range and pattern, where appropriate
	1.2 Population size	1.2.1 Population abundance and/or biomass, as appropriate
	1.3 Bycatch/discard	1.3.1 Bycatch/discard of selected fish species (target and non-target species, e.g. vulnerable species) in relation to population size
	1.4 Conservation status of fish	1.4.1 Conservation status of fish (CSF)
D3: Commercially exploited fish and shellfish	3.1 Level of pressure of the fishing activity	3.1.1 Fishing mortality (F) 3.1.2 Ratio between catch and biomass index (harvest ratio: HR))
	3.2 Reproductive capacity of the stock	3.2.1 Spawning Stock Biomass (SSB) 3.2.2 Biomass indices (CPUE)
	3.3 Population age and size distribution	3.3.1 Proportion of fish larger than the mean size of first sexual maturation (L_{mat}) 3.3.3 95% percentile of the fish length distribution observed in research vessel surveys (L95)
D4: Food webs	4.2 Proportion of selected species at the top of food webs	4.2.1 Large fish (by weight) (LFI)

considering those criteria referring to indicator application; and (III) evaluation of the theoretical basis, considering those criteria which concern data quality and the theoretical concept. It was assumed that these evaluation steps will help in identifying pros and cons of indicators more accurately.

The importance of each criterion was individually weighted, ranging from “essential” via “desirable” to “informative”. Indicator compliance was measured as “fully fitted” or “not fitted”, and beyond for the majority of the criteria a third intermediate level was used, “partially fitted”. To achieve a quantitative evaluation result the weightings assigned to each criterion were given scores of “essential” = 3, “desirable” = 2, “informative” =1 and compliance fits were scored with “fully fitted” = 1.0, “partially fitted” = 0.5 and “not fitted” = 0.0. The performance score of each indicator against each criterion is the product of these two values. In a next step summing across all criteria results in an overall score of indicator performance. Here criterion 1 (distinguishing state and pressure indicators) is not included in the calculation.

The experts involved were from the Thünen-Institutes of Sea Fisheries and Baltic Sea Fisheries in Germany. In order to ensure consistency between their ratings, guidelines were used for assessing each indicator against each criterion. Each expert could choose not to rate a certain indicator as a result of lack of detailed expertise. Each indicator was scored by at least 5 and at maximum 7 experts. To summarize indicator evaluation results of all experts, the mean of all expert scores of indicator compliance per criterion was calculated. Then by multiplying with the respective weighting factor a sum-product is calculated per indicator. Divided by the total possible amount of scores and multiplied by 100 the final result describes indicator performance as percentage.

$$\text{indicator score} = \frac{n\sum_{i=1}^{14} mean\ compliance\ score * \ importance\ score}{maximum\ possible\ indicator\ score} * 100$$

For the purpose of the present study the list of criteria, their weighting and compliance rating as developed by WGBIODIV were slightly adapted. Decisions were based on discussions of the whole group and were achieved as consensus. Table 1 in Annex I lists the criteria used, importance weightings and guidelines for assessing indicator compliance, corresponding to WGBIODIV.

The criteria are essentially intended to assess the performance of state indicators. Nevertheless the MSFD fish indicators also contain pressure indicators. Following the reasoning of WGBIODIV these indicators were treated differently. Their scoring against criteria related to sensitivity and responsiveness can be considered inappropriate since pressure indicators fulfill these aspects per definition wherefore they would score unfairly high (ICES 2013). Thus in a first step criterion 1 distinguished between state and pressure indicators. In the following only state indicators were rated against criteria 2-15, whereas pressure indicators were not applied for the only state relevant criteria 6, 8, 12 and 13.

2.2 Definition of indicator thresholds
In order to be able to interpret evaluation results objective benchmark thresholds were developed with a bootstrap method. In a randomized process 5000 performance scores for virtual indicators were simulated to get a likely frequency distribution of values. In a next step the benchmark threshold was statistically chosen as the 95% confidence level. According to WGBIODIV (ICES 2013) this benchmark score balances error rates of type I ("(rejection of a null hypothesis that is true, equivalent to accepting an indicator that actually performs poorly") and type II ("failure to reject a null hypothesis that is false, equivalent to "rejecting" an indicator that actually performs well"). Computing multiple virtual indicator frequency distributions and subsequent definition of indicator thresholds with the bootstrap method was done separately for state and pressure indicators as well as for the three evaluation aspects (I) overall indicator performance, (II) operationality, and (III) theoretical basis. According to the procedure of WGBIODIV each set of simulated criteria for calculated either for a virtual state indicator or a virtual pressure indicator (ICES 2013). In the case of a virtual state indicator the following formula was used:

$$\text{benchmark state indicator} = \frac{\sum importance\ score * compliance\ score}{\sum importance\ score}$$

In the case of a virtual pressure indicator a subset of the compliance score was set zero. Table 2 shows the resulting thresholds per evaluation aspect for state and pressure indicators respectively. Only when evaluating operationality state and pressure indicators do have the same thresholds, since all criteria relevant for assessing operationality address both state and pressure indicators.

Table 2: Indicator thresholds from simulations for each evaluation aspect; 95% percentile for the distribution of randomized indicator scores, distinguishing between state and pressure indicators.

Evaluation aspect	State indicator	Pressure indicator
Overall ind. performance	69 %	55 %
Operationality	76 %	76 %
Theoretical basis	71 %	55 %

3 Results

State and pressure indicators have to be treated differently when interpreting scoring results in respect to indicator performance, since by definition pressure indicators cannot score as high as state indicators, given that they are automatically set to zero against four criteria (ICES 2013). This is justified by WGBIODIV with the requirements of the MSFD to monitor progress towards GES, wherefore state indicators should be preferred (ICES 2013).

In the following the scoring results are described and interpreted in detail for each descriptor separately.

<u>Descriptor 1</u>

Indicator 1.2.1 on the assessment of population size (biomass or abundance) of non-commercial fish species shows high indicator performance. Nevertheless by taking a closer look at the detailed criteria scoring an important weakness can be identified (table 3.2): The indicator is not responsive to a specific pressure since it reacts to integrated effects of recruitment and mortality; recruitment is influenced by climate and habitat quality, mortality is influenced by fishing, predation pressure and eutrophication (HELCOM 2012).

Indicator 1.1.1/1.1.2 on distributional range and pattern of non-commercial fish species (hereinafter referred to as “indicator on fish distribution”) showed good indicator performance with similar scores as the indicator on fish population size. Nevertheless differences exist and the detailed criteria scores highlight that the indicator on fish distribution faces more limitations (table 3.1): Most notably this indicator was considered to be even less specific to pressures. The distribution and especially distributional patterns of fish populations can be expected to be extremely sensitive to multiple effects like the above mentioned ones and beyond. Furthermore fish distribution may be correlated to fish population size, wherefore it should be proved case-specifically, if the indicator shows redundancy to indicator 1.2.1. Moreover compared to the indicator on fish population size the fish distribution indicator showed poorer performance referring to early warning function as well to quantifiability. Although the indicator on fish distribution was considered established as national indicator it has to be made clear that it still needs further international development and validation.

The pressure indicator 1.3.1 on bycatch/discard of selected fish species (target and non-target species, e.g. vulnerable species) in relation to population size showed good overall indicator performance and a good theoretical basis but achieved only low scoring for its operationality. Concerning this matter an important aspect is the availability of data. Usually bycatch rates can be only be estimated for some commercially important species and even some of these lack proper estimates due to high variability according to season, nation and/or daily market prices. For non-commercial fish species, especially vulnerable species, not only bycatch monitoring needs improvement but also population sizes are often unknown. Generally discard rates are not known at all due to lacking possibilities for sufficient observation.

Indicator 1.4.1 on conservation status of fish (CSF) was scored low concerning overall indicator performance, although operationality and theoretical basis still scored high. This is due to the low scoring of the following reasons: Distinctness of the indicator scored low since the CSF is very sensitive towards the method of calculation including selection of species and reference period. Furthermore it is one problem of the CSF that usually scientific surveys started after commercial fisheries. Accordingly often there is no comparison possible to fish abundances in ecosystems not impacted by fisheries and vulnerable species, for which data is needed for calculation of CSF, are often not caught representatively in these surveys if they already declined. Therefore the indicator scored low according to relevance to management objectives. Specifity to pressures was considered to be weak as well since the CSF is a community indicator. As a consequence it might be difficult to entangle various explanations for low abundances of certain species since these can occur because of fisheries but other threats as well, e.g. pollution, introduction of non-indigenous species or climate. Thus pressure-state relationships referring to the conservation status of individual species might be difficult to detect by the indicator which is again hindering the implementation of appropriate management measures. In addition the underlying calculations are usually updated in large time intervals. This is also a reason for scoring low the indicators' early warning function. Finally the CSF is not required by the European Commission (EU-COM 2010) wherefore it scored low relating to its compliance of metric with MSFD indicator function.

<u>Descriptor 3</u>

Scoring results of the D3 indicators on commercial fish species indicate high overall indicator performance. Nevertheless detailed consideration of indicator scores per criterion highlights strength and weaknesses of each indicator, as is shown by the following tables 3.5- 3.10. Hence the indicators 3.1.1 (F), 3.2.1 (SSB) and 3.2.2 (CPUE) which are already established and used within stock assessments by the International Council for the Exploration of the Sea (ICES) achieve extraordinarily high scores within all three criterion categories (quality of underlying data, management, and conception). In contrast the indicators 3.1.2 (HR), 3.3.1 (L_{mat}) and 3.3.3 (L95) which are not established so far need for the necessary further development careful consideration of certain problems, which are highlighted by the corresponding criteria in question.

The indicator HR is a secondary indicator, proposed for the assessment of fishing mortality in the case of data limitations hindering the calculation of the preferred primary indicator F. Although being a pressure indicator it shows good overall performance. Major problems identified by Probst et al. 2014 were the facts that the indicator is trend based and as a consequence does not give a proper estimate of sustainability of the fishery according to the concept of maximum sustainable yield (MSY). Accordingly the theoretical basis scores low.

The indicators L_{mat} and L95 both score high concerning their theoretical basis, data availability, quantifiability, spatial coverage and cost-effectiveness. But consideration should be directed towards one of the indicators' major weakness relating to specificity to pressures. Ideally an indicator should reflect state change that is caused by a specific significant manageable pressure (ICES 2013). Both L_{mat} and L95 can be considered to respond not only to known pressures like fishing mortality but also to impacts of further state changes (e.g. recruitment). Therefore this criterion was scored to be only partially met. Especially the effect of varying recruitment has to be evaluated carefully. Since both indicators are relative metrics variably changing amounts of small fish in survey catches can influence the resulting calculated proportion of large fish significantly (Probst et al. 2012). This can interfere with the distinct interpretation of indicator trends. Accordingly the indicators also only partially meet the criterion relevance to management measures. The extent of any change in human activity required is unknown, only advice can be given on the required direction of action. Furthermore L_{mat} and L95 score low concerning another criterion of a good indicator which is the function of early warning. Both indicators were evaluated by the majority of experts to lack the ability to signal potential future state change and not before this is indicated by other MSFD indicators. Nevertheless the according very low scoring of this criterion did not lead to a bad overall performance, since the importance was weighted to be "informative" =1.

Descriptor 4

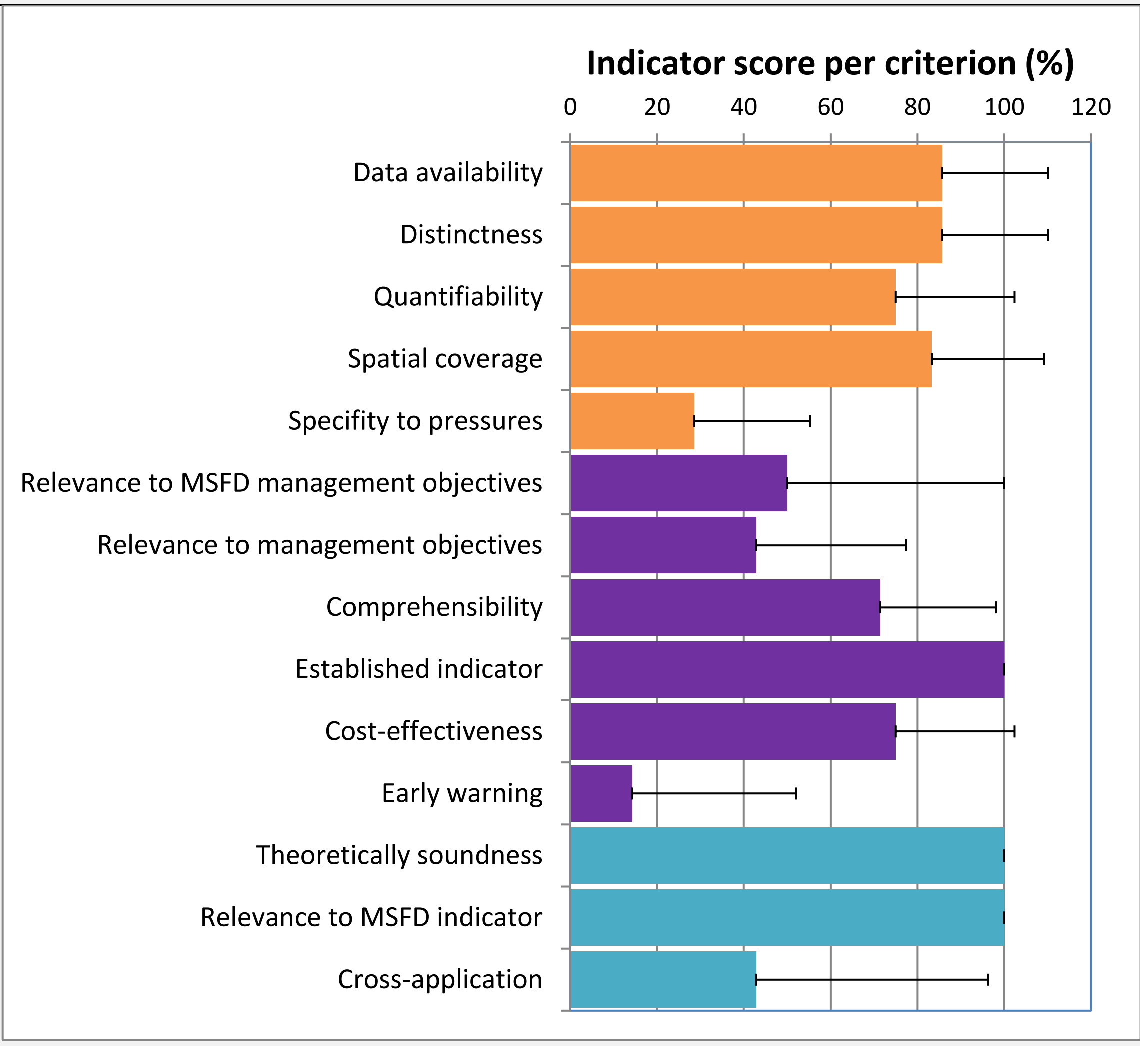
Indicator 4.2.1 (LFI) shows good overall performance and high scores for its theoretical basis and operationality as well (table 3.11). It was evaluated as being an established indicator but only regarding the North Sea, where the indicator was developed in the frame of developing OSPAR Ecological Quality Objectives (EcoQO) for the North Sea demersal fish community (Greenstreet et al. 2011). In the Baltic Sea the indicator is still under development in the frame of HELCOM. While the indicator is fully operational for the North Sea with eventual minor adaptions for local fish communities, the LFI faces major problems in the Baltic Sea where previous analyses showed that the indicator for the demersal fish community is dominated by one single species namely cod (Oesterwind et al. 2013). It has to be tested if and how the indicator can still fulfil its intended role as community indicator in the Baltic Sea or if it makes no sense at all to use it for the area in the frame of the MSFD. Beyond these issues the indicator generally scored low concerning its capability of early warning. This is due to the fact that the LFI was shown to respond to fishing pressure on a fish community with a time-lag of > 12 years (Greenstreet et al. 2011).

Acknowledgements

The research leading to these results is part of the BIO-C3 project and has received funding from BONUS, the joint Baltic Sea research and development programme (Art 185), funded jointly from the European Union's Seventh Programme for research, technological development and demonstration and from national funding institutions.

1.1.1/1.1.2 DISTRIBUTIONAL RANGE AND PATTERN, WHERE APPROPRIATE

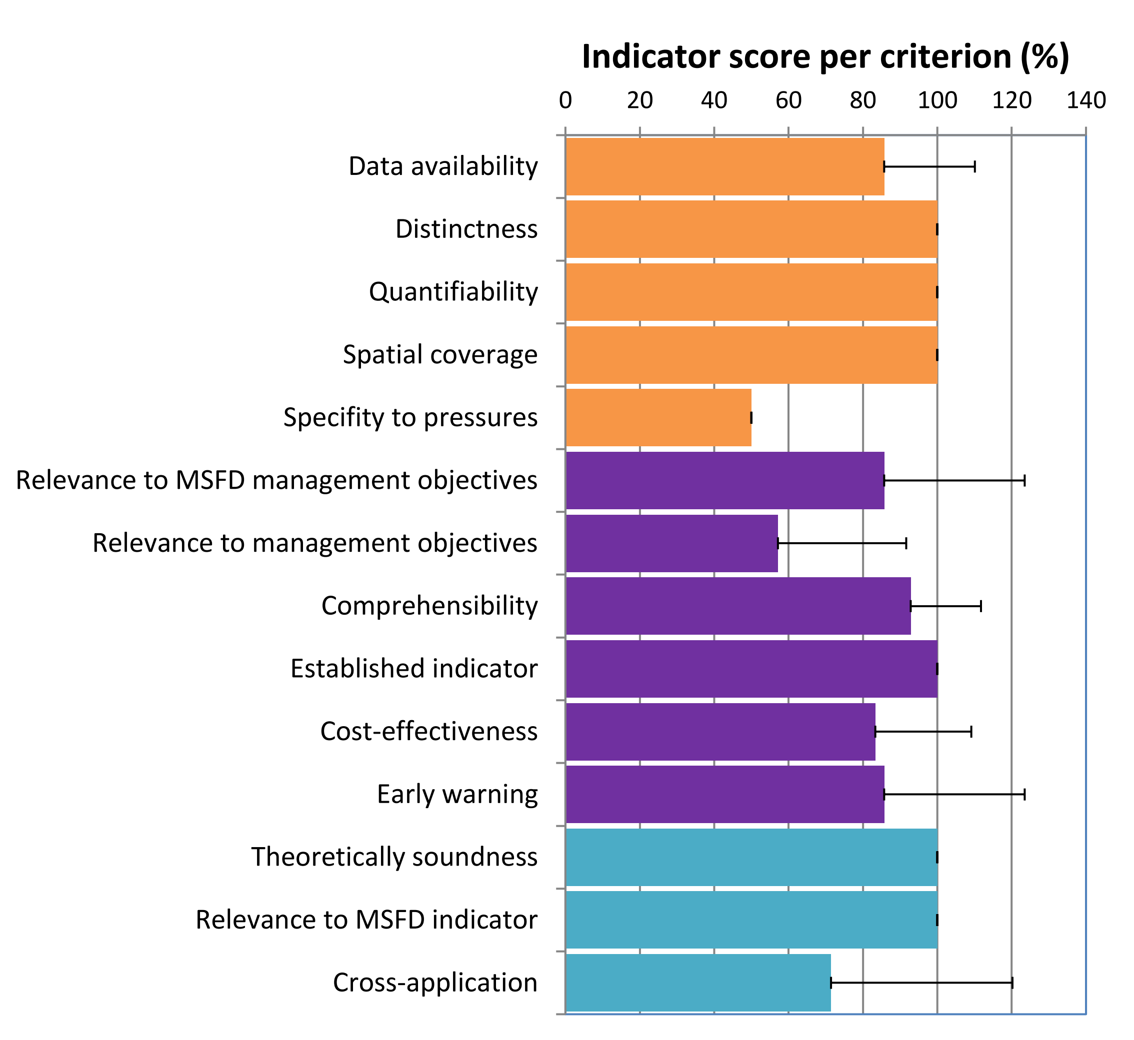
Table 3.1: Scoring results for combined MSFD indicators 1.1.1/1.1.2 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	86.48
Overall benchmark threshold (%)	69
Overall indicator performance	High
Operationality score (%)	91.48
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	89.68
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	High

1.2.1 POPULATION ABUNDANCE AND/OR BIOMASS, AS APPROPRIATE

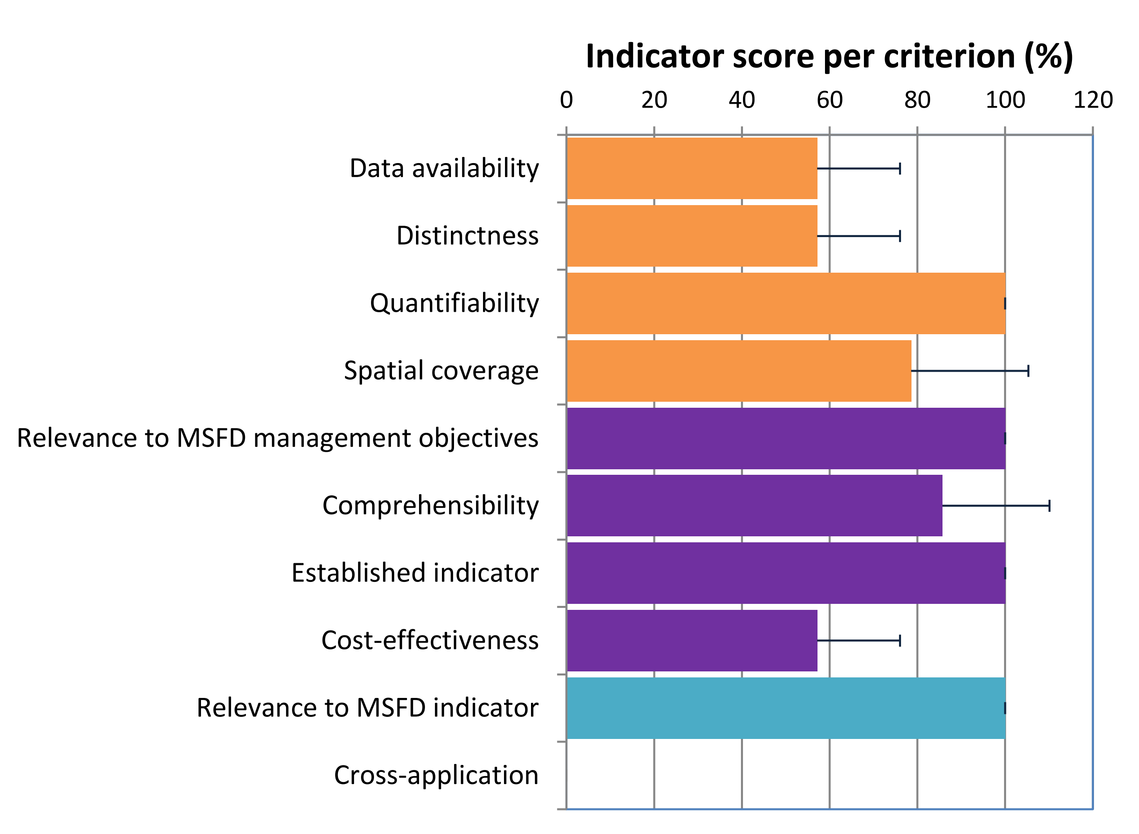
Table 3.2: Scoring results for MSFD indicator 1.2.1 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	86.48
Overall benchmark threshold (%)	69
Overall indicator performance	High
Operationality score (%)	91.48
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	89.68
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	High

1.3.1 BYCATCH/DISCARD OF SELECTED SPECIES (TARGET AND NON-TARGET SPECIES, E.G. VULNERABLE SPECIES) IN RELATION TO POPULATION SIZE

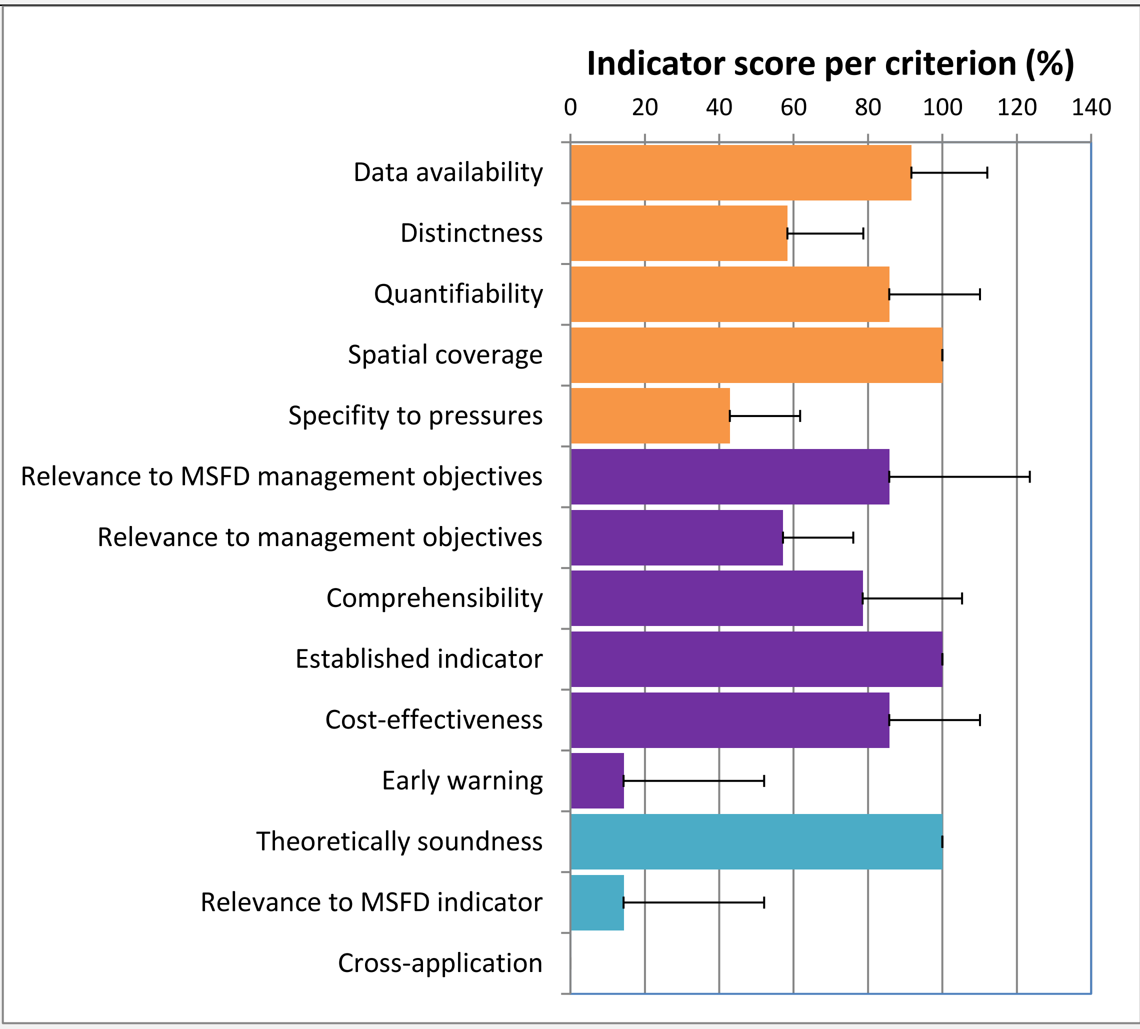
Table 3.3: Scoring results for the German MSFD indicator 1.3.1 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	55.07
Overall benchmark threshold (%)	55
Overall indicator performance	High
Operationality score (%)	68.05
Benchmark threshold operationality (%)	76
Operationality performance	Low
Theoretical basis score(%)	60.71
Benchmark threshold theoretical basis (%)	55
Theoretical basis performance	High

1.4.1 CONSERVATION STATUS OF FISH

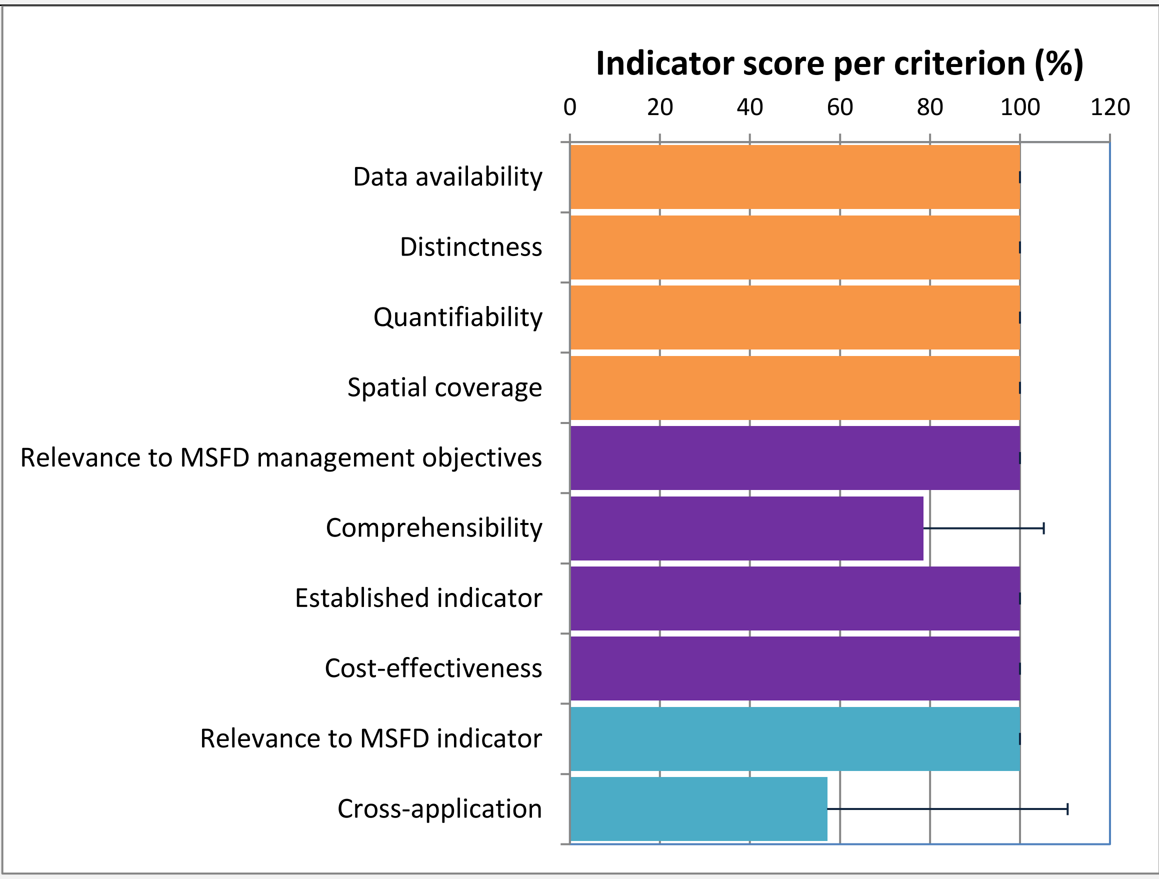
Table 3.4: Scoring results for the German MSFD indicator 1.4.1 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	65.09
Overall benchmark threshold (%)	69
Overall indicator performance	Low
Operationality score (%)	77.07
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	79.17
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	High

3.1.1 FISHING MORTALITY (F)

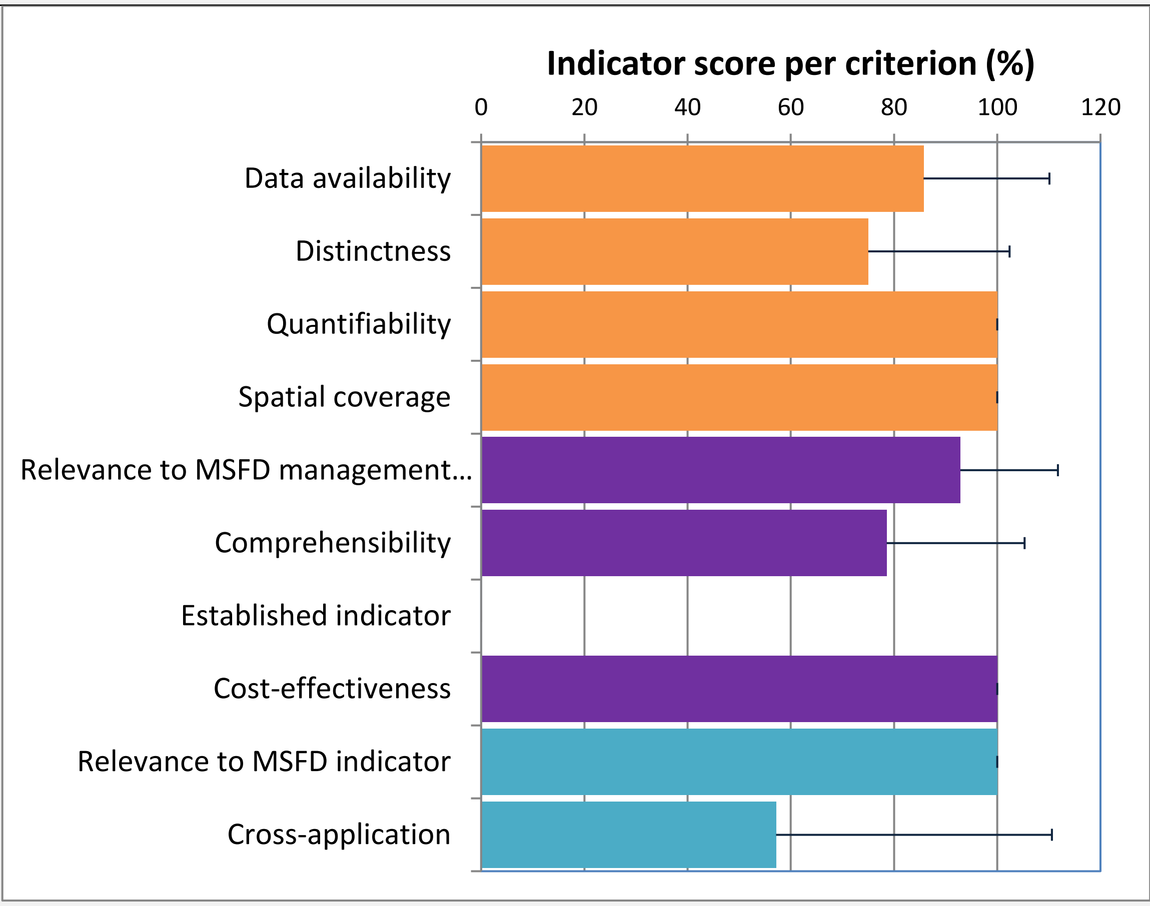
Table 3.5: Scoring results for the German MSFD indicator 3.1.1 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	70.74
Overall benchmark threshold (%)	55
Overall indicator performance	High
Operationality score (%)	95.49
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	71.03
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	High

3:1:2 RATIO BETWEEN CATCH AND BIOMASS INDEX (HR)

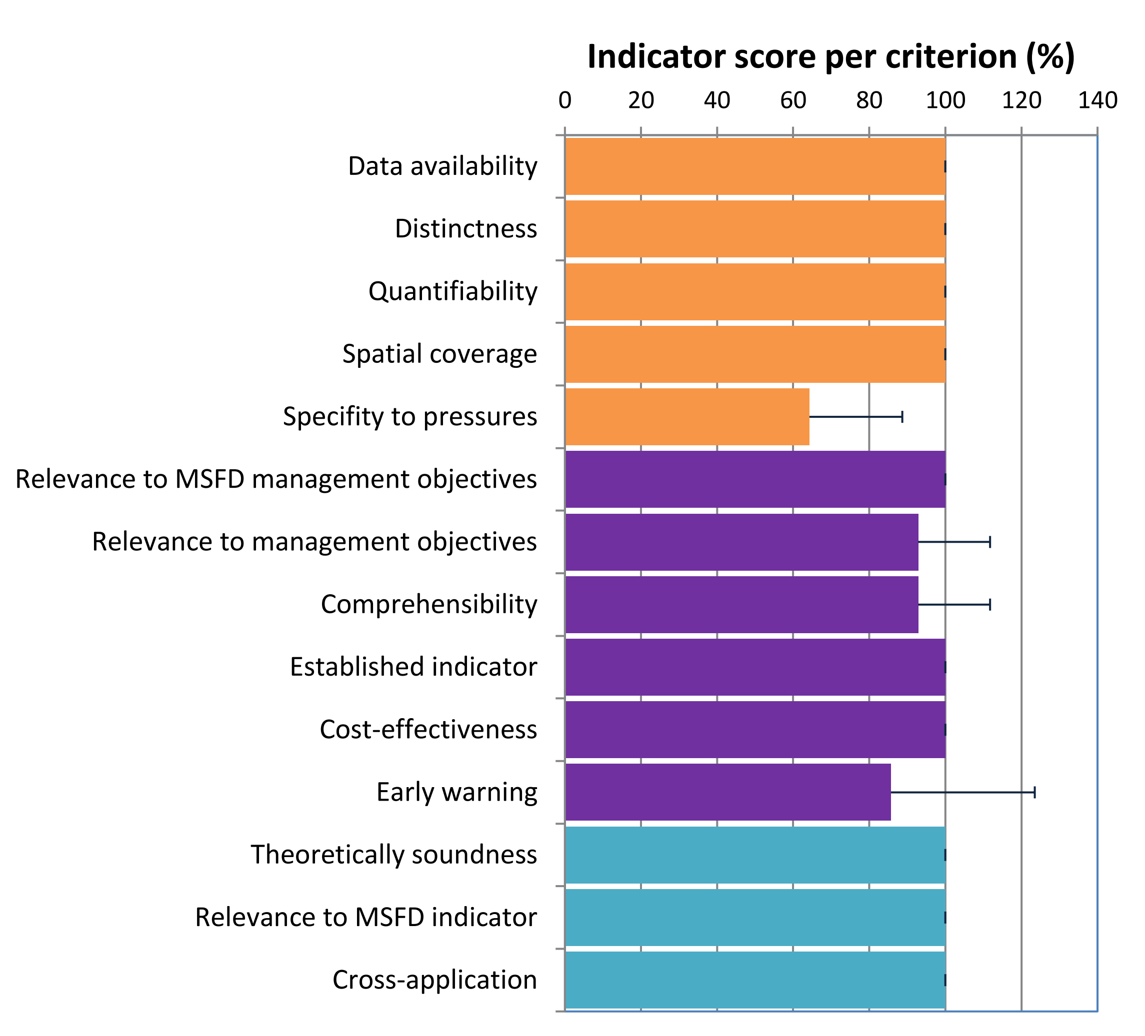
Table 3.6: Scoring results for the German MSFD indicator 3.1.2 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	60.02
Overall benchmark threshold (%)	55
Overall indicator performance	High
Operationality score (%)	78.01
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	54.96
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	Low

3.2.1 SPAWNING STOCK BIOMASS (SSB)

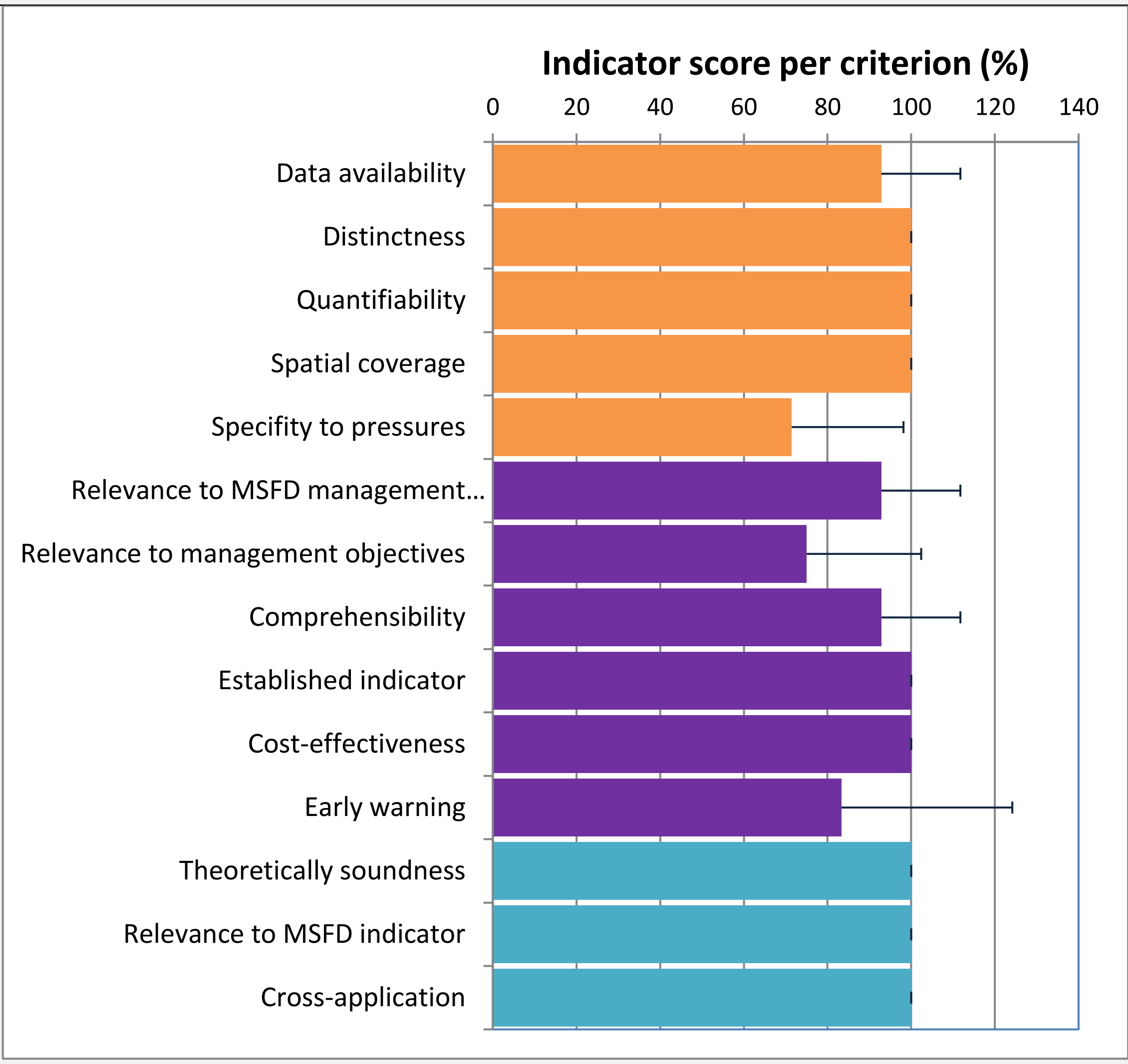
Table 3.7: Scoring results for MSFD indicator 3.2.1 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	95.39
Overall benchmark threshold (%)	69
Overall indicator performance	High
Operationality score (%)	100.00
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	93.65
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	High

3.2.2 BIOMASS INDICES (CPUE)

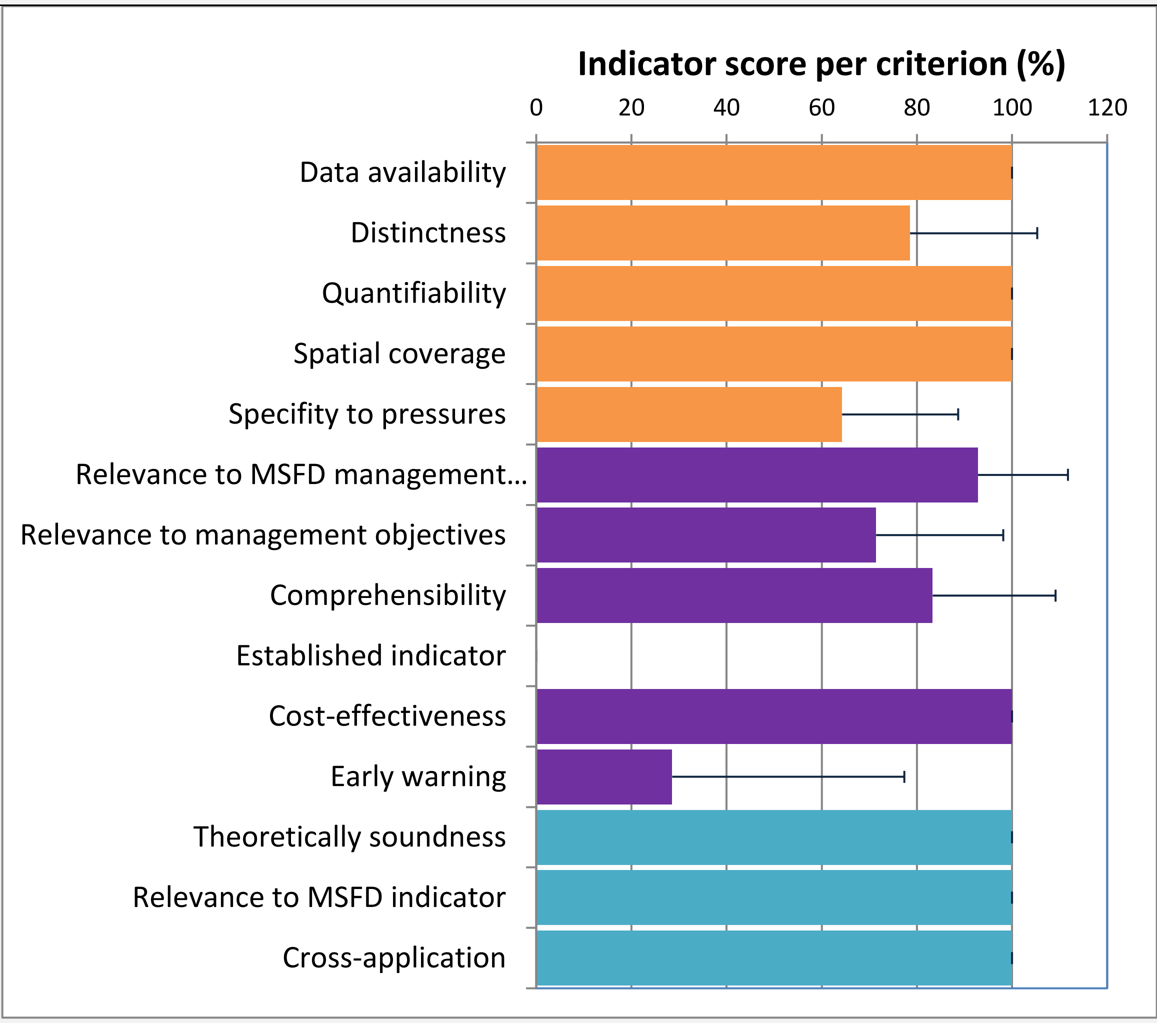
Table 3.8 Scoring results for MSFD indicator 3.2.2 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	93.70
Overall benchmark threshold (%)	69
Overall indicator performance	High
Operationality score (%)	98.12
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	94.05
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	High

3.3.1 PROPORTION OF FISH LARGER THAN THE MEAN SIZE OF FIRST SEXUAL MATURATION (L_{MAT})

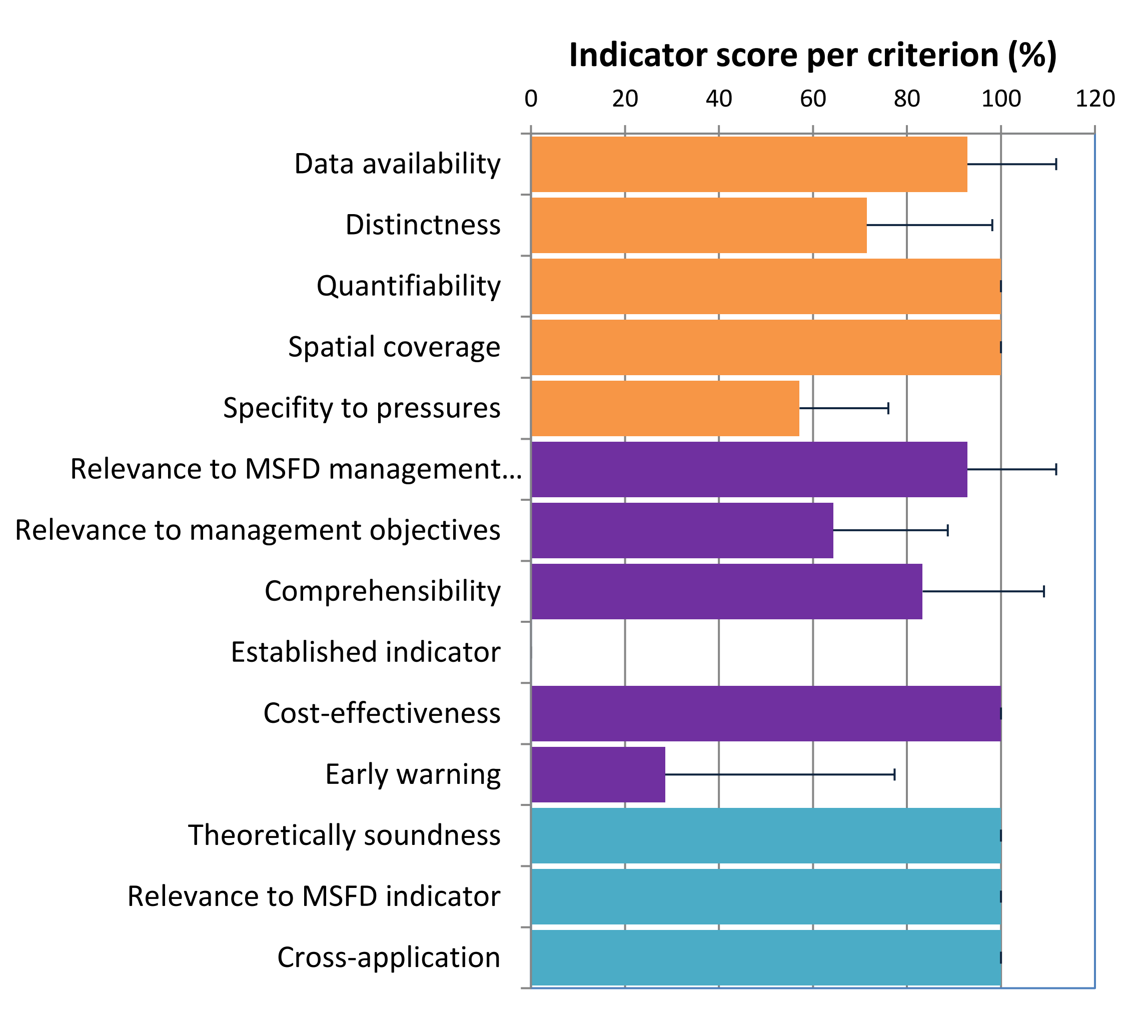
Table 3.9 Scoring results for MSFD indicator 3.3.1 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	82.87
Overall benchmark threshold (%)	69
Overall indicator performance	High
Operationality score (%)	85.34
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	77.65
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	High

3.3.3 95% PERCENTILE OF THE FISH LENGTH DISTRIBUTION OBSERVED IN RESEARCH VESSEL SURVEYS (L95)

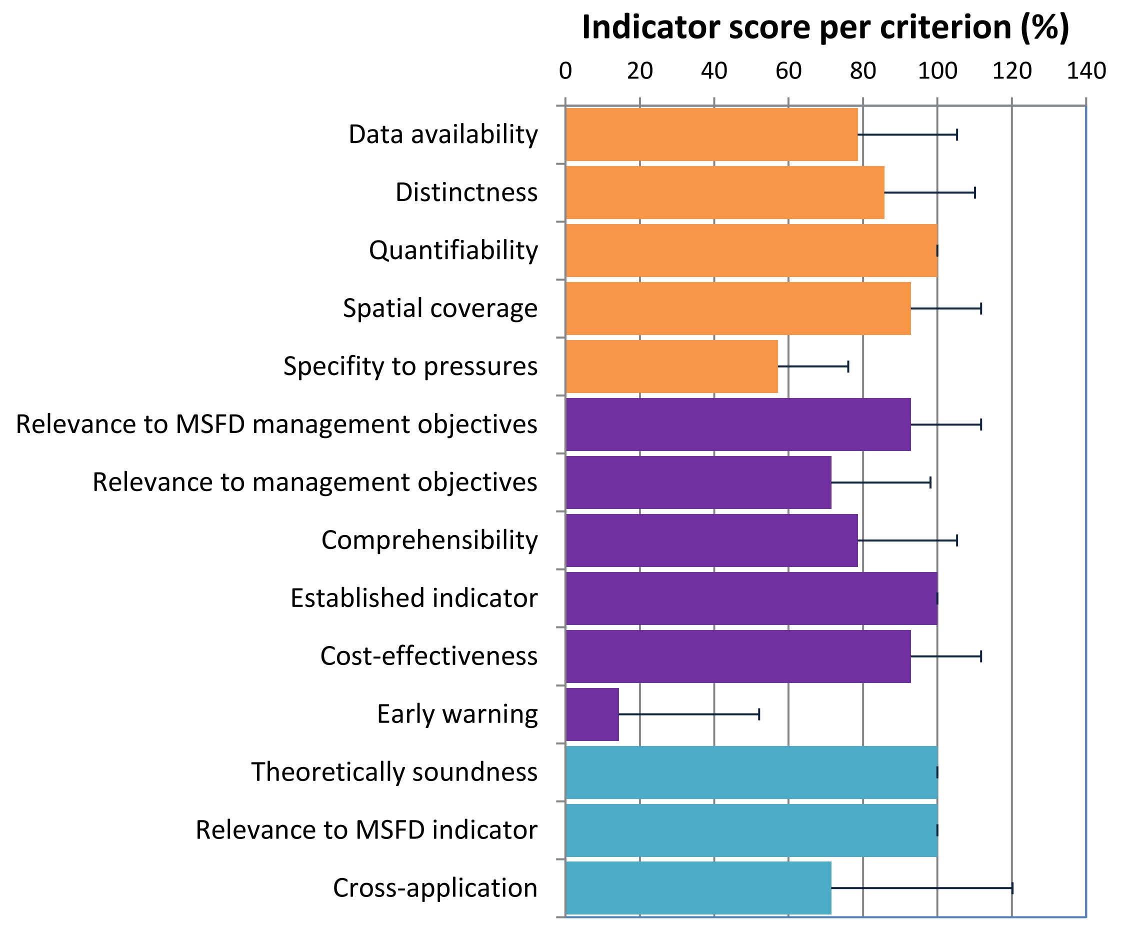
Table 3.10 Scoring results for MSFD indicator 3.3.3 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	80.34
Overall benchmark threshold (%)	69
Overall indicator performance	High
Operationality score (%)	83.08
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	75.26
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	High

4.2.1 LARGE FISH BY WEIGHT (LFI)

Table 3.11 Scoring results for MSFD indicator 4.2.1 (Probst et al. 2014). Colors indicate criterion category: orange = quality of underlying data; purple = management; blue = conception.



Total score (%)	83.64
Overall benchmark threshold (%)	69
Overall indicator performance	High
Operationality score (%)	87.90
Benchmark threshold operationality (%)	76
Operationality performance	High
Theoretical basis score(%)	87.30
Benchmark threshold theoretical basis (%)	71
Theoretical basis performance	High

Acknowledgements

This work resulted from the BONUS Bio-C^3 project and was supported by BONUS (Art 185), funded jointly by the EU and through the BMBF (Grant No. 03F0682B).

Literature

EU-COM (2008): Directive 2008/56/EC of the European parliament and of the council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive)

EU-COM (2010): Commission decision of 1 September 2010 on criteria and methodological standards on good environmental status of marine waters

Greenstreet, S.P.R., Rogers, S.I., Rice, J.C., Piet, G.J., Guirey, E.J., Fraser, H.M., Fryer, R.J., 2011: Development of the EcoQO for the North Sea fish community, ICES Journal of Marine Science, 68: 1-11.

HELCOM, 2012. Development of a set of core indicators: Interim report of the HELCOM CORESET project. PART B: Descriptions of the indicators. Balt. Sea Environ. Proc. No. 129 B.

ICES 2013. Report of the working group of biodiversity science (WGBIODIV), Copenhagen, Denmark.

ICES 2014. Recommendations for the assessment of MSFD Descriptor 3.

Oesterwind, D., Psuty, I., Pachur, M., Dorrien, C.von, Lejk, A., 2013: Proportion of large fish in the community, HELCOM Core Indicator Report. http://www.helcom.fi/Core%20Indicators/HELCOM-CoreIndicator-Proportion_of_large_fish_in_the_community.pdf

Probst, W.N., Rau, A., Diekmann, R., Dorrien, C. von, Seidel, H., Fock, H.O., Kraus, G., Stelzenmüller, V., 2014. Eine Thünen-Evaluierung von fisch- und fischereibezogenen Indikatoren der EU Meeresstrategie-Rahmenrichtlinie (MSRL). Hamburg: Johann Heinrich von Thünen-Institut, 106 p, Thünen Working Paper 25, DOI:10.3220/WP_25_2014.

Probst , W.N., Stelzenmüller, V., Fock, H.O., 2012. Using cross-correlations to assess the relationship between time-lagged pressure and state indicators: an exemplary analysis of North Sea fish population indicators. ICES Journal of Marine Science, 69(4): 670-681.

Annex I

Table 1: WGBIODIV criteria (ICES 2013), partly revised and used by the Thünen expert group to evaluate the performance of MSFD fish related indicators (Probst et al. 2014).

Criterion No.	Category	Criterion	Description of criterion	Importance Weighting	Importance Score A	Guidelines for Compliance Assessment Score B
1	Type of indicator	State or pressure	Is indicator a "pressure" indicator being used for want of an appropriate "state" indicator?			Fully met (1): indicator is a "state" indicator; Not met (0): indicator is actually a "pressure" indicator
2	Quality of underlying data	Data availability	Indicators must be supported by current or planned monitoring programs that provide the data necessary to derive the indicator. Ideal monitoring programs should have a time-series capable of supporting baselines and reference point setting. Data should be collected on multiple sequential occasions using consistent protocols, which account for spatial and temporal heterogeneity.	Essential	3	Fully met (1): long-term and ongoing data from which historic reference levels can be derived and past and future trends determined; Partially met (0.5): no baseline information, but ongoing monitoring; Not met (0): No monitoring.
3	Quality of underlying data	Distinctness	Metrics used to populate indicators should ideally be easily and accurately determined using technically feasible and quality assured methods.	Essential	3	Fully met (1): data and methods are technically feasible and quality assured in all aspects; Partially met (0.5): potential issues with quality assurance, or methods not widely adopted; Not met (0): metric is not tangible or doubtful.
4	Quality of underlying data	Quantifiability	Quantitative measurements are preferred over qualitative, categorical measurements, which in turn are preferred over expert opinions and professional judgments.	Desirable	2	Fully met (1): all data for the metric are quantitative; Partially met (0.5): data for metric are semi-quantitative or largely qualitative; Not met (0): metric is largely based on expert judgement.

Criterion No.	Category	Criterion	Description of criterion	Importance Weighting	Importance Score A	Guidelines for Compliance Assessment Score B
5	Quality of underlying data	Spatial coverage	Data should be derived from a relevant proportion of the MSFD subregion to which the metric will apply.	Essential	3	Fully met (1): spatially extensive monitoring is undertaken across the subregion; Partially met (0.5): monitoring does not cover the full subregion, but is considered adequate to assess status at subregional scale; Not met (0): monitoring is undertaken across a limited fraction of the subregion and considered inadequate to assess status at subregional scale.
6	Quality of underlying data	Specifity to pressures	The indicator reflects change in the state of an ecological component that is caused by a specific significant manageable pressure (e.g. fishing mortality, habitat destruction). The indicator should therefore respond sensitively to particular changes in a pressure.	Essential	3	IF CRITERION 1 IS SCORED 0 THEN THE SCORE MUST BE 0. Otherwise: Fully met (1): metric is responsive to a specific pressure and the pressure-state relationship is defined; Partially met (0.5): metric responds to several pressures and the pressure-state relationship is defined for at least one of these; Not met (0): no clear pressure-state relationship is evident.
7	Management	Relevance to MSFD management objectives	Clear targets (absolute values or trend directions) for the indicator can be specified that reflect management objectives.	Desirable	2	Fully met (1): an absolute or target or a target trend direction for the metric is set; Not met (0): targets or trends unknown.

Criterion No.	Category	Criterion	Description of criterion	Importance Weighting	Importance Score A	Guidelines for Compliance Assessment Score B
8	Management	Relevance to management objectives	Indicator links directly to management response. The relationship between activity and resulting ecological pressure on the ecological component is clearly understood.	Desirable	2	IF CRITERION 1 IS SCORED 0 THEN THE SCORE MUST BE 0. Otherwise: Fully met (1): Both pressure-state and activity-pressure relationships are well defined - one can advise on the direction AND extent of any change in human activity required; Partially met (0.5): only the pressure-state relationship is well defined - one can only advise of the direction of change in human activity required; Not met (0): no
9	Management	Comprehensibility	Indicators should ideally be understandable by policy-makers and other non-scientists (e.g. stakeholders) alike, and the consequences of variation in the indicator should be easy to communicate.	Informative	1	Fully met (1): the metric is easy to understand and communicate; Partially met (0.5): a more complex and difficult to understand metric, but one for which the meaning of change in the metric value is easy to communicate; Not met (0): the metric is neither easy to understand or communicable.
10	Management	Established indicator	Established indicators are preferred over novel indicators that perform the same role.	Desirable	2	Fully met (1): the indicator is established and used in international and/or national policy frameworks; Not met (0): the indicator has not previously been used in a management framework.

Criterion No.	Category	Criterion	Description of criterion	Importance Weighting	Importance Score A	Guidelines for Compliance Assessment Score B
11	Management	Cost-effectiveness	Sampling, measuring, processing, analyzing indicator data, and reporting assessment outcomes, should make effective use of limited financial resources.	Desirable	2	Fully met (1): little additional costs (no additional sampling is needed); Partially met (0.5): new sampling on already existing programs is required; Not met (0): new sampling on new monitoring programs is necessary.
12	Management	Early warning	Indicators that signal potential future change in an ecosystem attribute before actual harm is indicated by other MSFD indicators should be preferred. These could facilitate preventive management, which could be less costly than restorative management.	Informative	1	IF CRITERION 1 IS SCORED 0 THEN THE SCORE MUST BE 0. Otherwise: Fully met (1): indicator provides early warning because of its high sensitivity to a pressure with short response time; Not met (0): relatively insensitive indicator that is
13	Conception	Theoretically soundness	Scientific, peer-reviewed findings should underpin the assertion that the metric provides a true representation of variation in the ecosystem attribute in question.	Desirable	2	IF CRITERION 1 IS SCORED 0 THEN THE SCORE MUST BE 0. Otherwise: Fully met (1): peer-reviewed literature including description of indicator, target and analysis of important indicator criteria (like specifity, sensitivity, reactivity); Partially met (0.5): documented but not peer-reviewed; Not met (0): not documented, or peer-reviewed literature is contradictory.

Criterion No.	Category	Criterion	Description of criterion	Importance Weighting	Importance Score A	Guidelines for Compliance Assessment Score B
14	Conception	Relevance to MSFD indicator	For D1 and D3, metrics should fit the indicator function stated in the 2010 MSFD Decision document. This requirement can be relaxed for D4 indicators because the Decision document stipulates the need for indicator development in respect of this Descriptor (but any newly proposed D4 indicators must still fulfil the overall goals stated for D4).	Essential	3	Fully met (1): metric complies with indicator function; Not met (0): metric does not comply with indicator function.
15	Conception	Cross-application	Metrics that are applicable to more than one MSFD indicator are preferable.	Desirable	2	Fully met (1): metric is applicable across several MSFD indicators; Not met (0): no cross-application.